SOCIALLY MOTIVATED VOCAL LEARNING

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Supriya Syal

August 2011

SOCIALLY MOTIVATED VOCAL LEARNING

Supriya Syal, Ph. D.

Cornell University 2011

Language is social because the motivation to communicate is contingent on the presence of social conspecifics. This thesis studies vocal learning in two social species, humans (*Homo sapiens sapiens*) and an oscine songbird, the zebra finch (*Taeniopygia guttata)*. Using both behavioral and neuroanatomical data, Chapter 2 of this thesis argues that both the evolution and the development of vocal learning were triggered by the evolution of gated links between sub-cortical social motivation and cortical learning devices. Chapters 3 and 4 study human vocal learning in the socially motivated contexts in which human infants typically learn language. Chapter 3 studies early phonological learning within a social reinforcement paradigm wherein specific features of the interaction between mothers and infants guide the learning and production of new distributional information. Chapter 4 looks at the effects of changing maternal motivations on maternal behavior, and their subsequent facilitatory effects on infant vocal learning. Chapter 5 suggests that motivation derived from early interactions with social conspecifics drives vocal development in young zebra finches. Overall, this thesis provides evidence that vocal learning is shaped by social interaction and argues for a consideration social motivation as a control parameter in vocal learning.

BIOGRAPHICAL SKETCH

Supriya was born in India. When she was 6, she was told that when she grows up, she must go to the US and become a doctor. This has been her attempt to do so.

ACKNOWLEDGEMENTS

Barb Finlay: You Rock.

David Field, David Pizarro, Elizabeth Adkins Regan, Morten Christiansen, Tom Gilovich, and Harry Segal: You Rock too, albeit a bit lesser than Barb. But then again, don't we all?

Pam Cunningham, Julie Simmons-Lynch and Tim VanDeusen: You make it better.

Satyaki Prasad, Catalina Iricinschi, Mike Hove, Sarah Anderson, Lukas Kroc, Sudeep Mandal, Praveen Gowtham and Radha Desai: Thanks to you, I will never forget the importance of chicken curry, tarot reading, drums 'n bass jam, cursing-on-statistics, off-road driving, ice-cream, metaphysical fencing, and how-to-be-someone's-best-friend-on-Skype, in life, or in science.

KFC: You have the same number of syllables as Ph.D. Thank you for introducing spicy chicken wings in my third year – it was when I needed them the most.

Kelly: Hurray for you!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Language is strikingly social. While it is possible to have social behavior that is not linguistic, the converse is not true. Language is learned, perceived and produced within the fabric of social interaction. My work argues that language remains embedded in sociality because the *motivation*[1] to communicate exists only within a social context. This thesis takes a comparative and developmental approach to argue that sub-cortical motivational systems form an integral part of the "language circuit" and that the presence or absence of such motivational reward systems, and their structural and functional linkages with conventional language and social networks, are the critical determinants of the evolution and development of language in a given species.

My primary research goal is to study the role of social motivation in language learning. I am interested in studying language within the rich social context of the human ecological niche to better understand the underlying functional architecture that supports both its learning and usage. To that end, this dissertation is focused on the role of socially derived motivation through infant-caregiver interactions in the development of vocal communication in humans, and a birdsong model of human language, zebra finches.

Social motivation in the evolution and development of language

Learned vocal behavior is a feature that is found in few and far between but always social/gregarious species. The central thesis of my work is that vocal learning is grounded in motivated social behavior, and such functions have implications for the evolution and

[1] Etymologically, the Latin word *movere*, “to move”, meaning "that which inwardly moves a person to behave a certain way"

development of underlying neural mechanisms. In *Chapter 2*, we explore advances in the understanding of the social behavior network (Newman, 1999) and its evolutionary variations (Goodson, 2005) highlighting some key findings in vocal learning across species. In organisms that learn a complex language, social context seems to be necessary for this process (Fromkin et al., 1974; Lane, 1976; Marler, 1970; Thorpe, 1958; Kuhl, 2007). In organisms that don't specifically need to learn their species-typical vocalization (Owren et al., 1992; Seyfarth & Cheney, 1986; Winter et al., 1973), social context still modifies and modulates the nature and the content of such vocalizations (Roush & Snowdon, 1994, 1999; Seyfarth & Cheney, 1997). A central argument here is that organisms are *motivated* to communicate in a social setting that contains conspecifics.

In the cross species analysis of complex vocal learning, birdsong has often been used as a model for human language. While both song learning strategies and the final song repertoire vary greatly between different songbird species, they nonetheless provide an excellent system to study the evolution of communication systems (Beecher and Brenowitz, 2005). Perhaps the most striking parallel between human language and birdsong is the importance of vocal learning (production and perception) in the development and maintenance of species typical vocal behavior (Doupe & Kuhl, 1999; Goldstein et al., 2003). The parallels in comparative neuroanatomy, critical periods, mechanisms of learning, lateralization and so on are also striking (Seidenberg & Zevin, 2006).

Recent work in comparative developmental neuroanatomy has pointed at homologies in the embryological origins of human amygdala/basal forebrain limbic regions and avian song centers (Medina, 2007; Striedter, 1997; Puelles et al., 2000; Puelles 2001; Martinez-Garcia, Martinez-Marcos, & Lanuza, 2002; Martinez-Garcia, Novejarque, & Lanuza, 2007). These data

suggest that instead of being assigned a neocortical fate, bird song nuclei are more soundly positioned in regions that typify socio-motivational learning and regulation in mammals.

While language has been reliably shown to be dependent on the cortex in adult humans, this is the crystallized product of a lifelong learning process; it is only in the development of vocal communication that one can study the essential changes in the evolution of the brain that produced language. We propose that the *motivation to learn to communicate* forms an integral part of the development of language in humans. We present behavioral data from humans that underscore the role of social motivation in language learning and make specific proposals about changes in neuroanatomical circuitry that might have been used, over evolutionary time, to link communication with motivation and reward in human infants. Further, we hypothesize the development of a neuroanatomical link that constitutes a socio-motivational "gate" necessary and sufficient for the evolution of socially motivated vocal learning.

Social motivation in human language learning

My work with humans attempts to situate the development of specific features of early vocal learning in the socially motivated settings in which these are species-typically learnt. The role of motivation or affect during communicative interactions between human mothers and infants has long been appreciated but puzzlingly remains rather under-studied. Factors such as the prosody of caregivers' early speech to infants, along with facial movements, form non-linguistic affective behaviours that "attract and sustain the infants' attention to what will become a linguistically important stream of cues" (Locke, 1993). Mothers provide infants with highly structured information that is made more salient through its intrinsic links with social motivation. For example, mothers commonly and reliably use predictable prosodic contours to modulate

infant attention (Fernald, 1992; Kuhl et al., 2005). They engage in contingent turn-taking vocal interactions with their infants that facilitate vocal development (Goldstein et al., 2003). Finally, caregivers are known to use variation sets – sentences with partially overlapping syntactical structure – to aid word learning (Onnis et al., 2008). All of these structural aids to language learning are presented to the infant in a characteristically social environment that is replete with positive feedback!

Chapter 3

Recent work has highlighted the role of contingent responsiveness from conspecifics in shaping adult-like vocal forms from the immature vocalizations of infant learners in both songbirds (West and King, 1988) and human babies (Goldstein et al., 2003). Female brown-headed cowbirds facilitate the transition to mature song (*Molothrus ater*) by selectively responding to more advanced plastic song vocalizations (West and King, 1988). Similarly, human caregivers drive the development of adult-like speech through both silent and vocal reinforcement that is contingent on infant vocalizations (Goldstein et al., 2003; Goldstein & Schwade, 2008). The latter of these studies found that specific forms of contingent vocal feedback from caregivers cause infants to produce specific and corresponding advances in their own vocalizations. Infants do not imitate the phonemic surface features of the sound itself but extract the underlying phonological pattern and generalize this learning to their own novel vocal productions (Goldstein & Schwade, 2008a)

In *Chapter 3 (Goldstein, Syal, and Schwade)*[2] we tested infant ability to learn novel phonological forms from socially motivating contingent input that different in its variability. The rationale for assessing the ability to learn from variable input derives from studies of statistical

[2] First author M. H. Goldstein

learning that show that infants are able to learn the regularities of speech from statistically distributed input in both artificial (Saffran, Aslinn and Newport, 1996) and natural languages (Pelucchi, Hay, and Saffran, 2009). Further, both infants and adults able to generalize this learning better if the training exemplars are more variable (Gomez, 2002) and in at least in adults, it has been suggested that higher level abstraction of embedded information is aided by greater variability (Onnis et al., 2004)

Within the domain of phonological learning, it is known that by 9 months of age, infants have learnt to use phonotactic regularities to parse fluent speech (Mattys and Jusczyk 2001) Additionally, infants' judgment of phonetic categories is modulated by variability in the input (Maye, Werker and Gerken, 2002) and they use extract invariant underlying rules to discriminate phonetic information (Maye, Weiss, and Aslin, 2008)

If these mechanisms of learning from distributional input play a role in phonological development, infants who were presented with a large number of different exemplars in a contingent social interaction would be able to learn and generalize the rule and produce novel/non-imitative vocalizations that are of the new phonological form. In a naturalistic social interaction, we cued caregivers to present infants with a variable number of exemplars of a novel phonological form. Additionally, we controlled the contingency of the interaction whereby one group of caregivers' delivery of the exemplar(s) was contingent on infant vocalization, while the other's was yoked to match stimulation but not contingency levels. We found that both contingency and phonetic variability were necessary for infant learning of the new phonological form. Infants were also able to extract and abstract the underlying structure of contingently presented variable input produce new sound sequences that obeyed the regularities caregivers' speech. Infants who received non-contingent or invariant input did not learn. Thus these findings

serve as evidence for the role of social motivation in infants' learning of language-specific features of their ambient language.

Chapter 4

Within the ecological framework of vocal learning, the parent infant dyad constitutes a socially distributed system of learning (Goldstein & Schwade 2008b). Thus it stands to reason that changes in either parent or infant context will affect vocal learning, possibly through changes in their motivations. In *Chapter 4*, we looked at the effects of changing maternal motivations and their subsequent effects on infant vocal learning. A very pervasive change in infants' ability to interface with the environment is brought on by the onset of independent locomotion during the first year of life. It has been argued that locomotion represents an activity dependent developmental process (Thelen & Smith, 1994; Adolph, 1997) that catalyses developmental changes in multiple domains such as social communication, perception and affect (Piaget, 1954). The dynamic of parent-child interaction also changes, with both infants and parents undergoing behavioral changes after the onset of independent locomotion such as increased interactive play, increased social referencing (Campos, Kermoian, & Zumbahlen, 1992; Campos et al., 2000) and increased use of specific affective contours in verbal communication (Campos, Kermoian, & Zumbahlen, 1992). Recent data has also indicated that newly locomoting infants use social referencing while making decisions about motor actions, especially when faced by uncertainty in terms of locomotor capacity (Tamis-LeMonda, 2008).

Do changes in the social environment, that result from direct changes in locomotor capacity, impact infant vocal learning in any way? One study found that infants who were able to locomote were different from infants who couldn't locomote in terms of their social behavior (Gustafson, 1984). Further, locomoting infants produced a higher number of vocalizations than

non-locomoting infants. In order to control for developmental effects, infants who were not able to locomote were placed in a walker such that they could move around on their own. It was observed that these infants quickly changed their behavior to resemble that of infants who could locomote on their own. These newly locomoting infants produced a higher number of vocalizations and directed and greater looks and smiles at caregivers, indicating a change in their social motivation. However, while the total number of vocalizations was counted, the acoustic parameters of these vocalizations were not studied. Further, changes in caregiver behavior as the infant context changed from non-locomotor to locomotor, were not assessed. Previous experiments in our lab have found that changes in social contingency can impact infants' production of advanced vocalizations (Goldstein et al., 2003; Goldstein and Schwade, 2008a). Are caregivers motivated to change the manner in which they interact with their infants once they start locomoting?

We hypothesized that caregivers are sensitive to their infants' ability to locomote and manipulations of infant locomotor status will change parental behavior. Parents of newly locomoting infants are motivated to be more vigilant and as a result engage in more contingent social interaction, providing their infants with greater vocal learning opportunities. This in turn should facilitate vocal development in locomoting infants.

We find that over developmental time, the ability to locomote is related to changes in the acoustic form of prelinguistic vocalizations. The vocal repertoires of infants who are able to locomote are characterized by increased proportions of more advanced sounds, such as canonical syllables. In a real time manipulation, infants increased the proportion of the most advanced sounds in their repertoire as their locomotor status was changed from immobile to mobile. This effect on infant vocal behavior was mediated by changes in the dynamics of mother-infant

interaction. Increased maternal vigilance led to increased social interaction between the mother and the infant. This change in caregiver behavior in turn creates more opportunities to learn which infants capitalized on, thus propelling themselves forward on the trajectory of vocal development.

Social motivation in birdsong learning

My work with zebra finches (*Chapter 5)* seeks to explore the role of reward-based motivation in vocal learning in a more targeted manner than is possible with human infants. As aforementioned, birdsong and human speech have long been considered parallel systems in the study of vocal learning (Doupe and Kuhl, 1999; Goldstein et al., 2003). A prediction of the work presented in *Chapter 2* would be a specific association between species-typical vocal learning and activation of reward circuitry. Food reward is among the most reinforcing motivators that exist across species (Wise, 2006). In songbirds that learn the song of a specific tutor, is there an association between food and song?

Zebra finches are monogamous pair-bonding birds that display biparental care of young, with individual families ranging from 4-6 members. They live in large multi-family colonies (often over 100 birds). Despite this large proliferation of potential song tutors, juvenile males tend to learn the song of their fathers. While various factors affecting tutor choice such as morphology (Mann et al. 1991), song characteristics (Clayton, 1987) including female-directedness of song (Roper & Zann, 2006), proximity (Mann & Slater, 1995; Roper & Zann 2006) and relatedness (Roper & Zann 2006; Williams, 1990) have been investigated; the direct influence of motivation to learn from a given tutor has not been studied.

We reasoned that specific features of social interactions surrounding food between

juvenile males and their fathers are likely to endow his song with motivational significance. To assay for these, we made high-resolution behavioral observations of father-infant interactions in the period immediately after fledging and before the end of the sensitive phase of song learning. We focused on specific vocal behaviors that the father engaged in as he fed the young

Our research shows that juveniles beg for food preferentially from the father. The father's responses to begging include feeding and song. Between day 25 and day 30, the proportion of responses that include song increases dramatically. Co-occurrences between begging and song cannot be explained by chance associations or by increases in singing between the two ages. Thus the father associates feeding with adult male song over the course of early vocal development in these birds. The proportion of begging bouts that receive song rises as juveniles grow older and enter the sensitive period of song learning. Further, the change in proportion paternal responses that were comprised of male song predicts future song outcome measures. This indicates that early social interactions with the father create opportunities for the song to gain motivational power. Thus young male zebra finches are motivated to learn the song of their fathers because the father is the one that feeds and maintains the greatest amount of social interactivity with them. These data provide evidence that socially situated motivation drives the learning of species-typical vocal forms in zebra finches.

Overall Significance

My theoretical and empirical work promotes an interactionist view of development wherein new motor, cognitive and social learning serve as precursors to the emergence of language, within the framework of socially motivated vocal learning. I attempt to return attention to language in its ecological context, acquired in the early development of an obligatorily social,

gregarious, and altruistic species where the motivation to learn to communicate with conspecifics drives both its ontogeny and phylogeny.

REFERENCES

Beecher, M. D. & Brenowitz, E. A. (2005) Functional aspects of song learning in birds. *Trends in Ecology & Evolution*, 20, 143-149.

Campos, J. J., Kermoian, R., & Zumbahlen, M. R. (1992). Socioemotional transformations in the family system following infant crawling onset. In N. Eisenberg & R. A. Fabes (Eds.), *Emotion and its regulation in early development* (pp. 25-40). San Francisco: Jossey-Bass

Campos, Joseph, J., Anderson, D. I., Barbu-Roth, M. A., Hubbard, E. M., Hertenstein, M. J., Witherington, D. (2000). Travel Broadens the Mind. *Infancy*, Vol 1 (2), 149-219.

Clayton, N. S. (1987). Song tutor choice in zebra finches. *Anim. Behav*., 35, 714-721.

Doupe, A. J. and Kuhl, P. K. (1999). Birdsong and human speech: common themes and mechanisms. *Ann. Rev. Neurosci.*, 22, 567-631.

Fernald, A. (1992a). Maternal vocalizations to infants as biologically relevant signals: An evolutionary perspective. In J.H. Barkow, L. Cosmides, and J. Tooby (Eds.). *The adapted mind: Evolutionary psychology and the generation of culture.* Oxford: Oxford University Press.

Fromkin, V., Krashen, S., Curtis, S., Rigler, D., Rigler, M. (1974). The development of language in Genie: a case of language acquisition beyond the "critical period". *Brain Lang.,* 1, 81-107

Goldstein, M. H., King, A. P., and West, M. J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proc. Nat. Acad. Sci.,* 100(13), 8030-8035.

Goldstein, M. H., and Schwade, J. A. (2008a). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychol. Sci.,* 19(5), 515-523

Goldstein, M. H., & Schwade, J. A. (2008b) From birds to words: Perception of structure in social interactions guides vocal development and language learning. In M. S. Blumberg, J. H. Freeman, & S. R. Robinson (Eds.), *The Oxford Handbook of Developmental and Comparative Neuroscience*. Oxford University Press.

Gómez, R. Variability and detection of invariant structure (2002), *Psychol. Sci.* 13, 431-436.

Goodson, J. L. (2005). The vertebrate social behavior network: evolutionary themes and variations. *Hormones and Behavior, 48*, 11-22.

Gustafson, G.E. (1984). Effects of the ability to locomote on infants' social and exploratory behaviors: An experimental study. *Developmental Psychology*, 20(3), 397-405.

Kuhl, P. K., Coffey-Corina, S., Padden, D. and Dawson, G. (2005). Links between social and linguistic processing of speech in preschool children with autism: Behavioral and electrophysiological evidence. *Dev. Sci*., 8,1-12.

Kuhl PK. 2007. Is speech learning 'gated' by the social brain? *Dev. Sci*. 10: 110–20.

Lane, H. L. (1976). *The Wild Boy of Aveyron.* Harvard Univ. Press, Cambridge, MA.

Locke J. (1993). *The Path to Spoken Language*. Harvard Univ. Press, Cambridge, MA.

Mann, N. I., Slater, P. J. B., Eales, L. A., and Richards, C. (1991) The influence of visual stimuli on song tutor choice in the zebra finch, Taeniopygia guttata, *Animal Behaviour*, 42(2): 285-293

Mann, N. I., and Slater, P. J. B. (1995), Song tutor choice by zebra finches in aviaries, *Animal Behaviour*, Volume 49, Issue 3, March 1995, Pages 811-820

Martinez-Garcia, F., Martinez-Marcos, A., and Lanuza, E. (2002). "The pallial amygdala of amniote vertebrates: evolution of the concept, evolution of the structure." *Brain Res Bull* 57(3-4): 463-9.

Martinez-Garcia, F., Novejarque, A., and Lanuza, E. (2007). Evolution of the Amygdala in Vertebrates. In Jon H. Kaas ed. *Evolution of Nervous Systems*. Academic Press, New York. pp: 255-334.

Mattys, S.L. & Jusczyk, P.W. (2001). Phonotactic cues for segmentation of fluent speech by infants. *Cognition*, 78, 91-121.

Maye, J., Weiss D. J., and Aslin R. N. (2008) Statistical phonetic learning in infants: facilitation and feature generalization. *Dev Sci*. 11(1):122-34.

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can

affect phonetic discrimination. *Cognition*, 82 (3), B101-B111.

Medina, L (2007). Do Birds and Reptiles Posess Homologues of Mammalian Visual, Somatosensory and Motor Cortices? In Jon H. Kaas ed. *Evolution of Nervous Systems*. Academic Press, New York. Vol 2 pp: 163-194.

Newman, S. W. (1999). The medial extended amygdala in male reproductive behavior. A node in the mammalian social behavior network. *Annals of the New York Academy of Sciences, 877*, 242-257.

Onnis, L., Monaghan, P., Christiansen, M.H., & Chater, N. (2004). Variability is the spice of learning, and a crucial ingredient for detecting and generalising nonadjacent dependencies. *Proceedings of the 26th Annual Conference of the Cognitive Science Society.*

Onnis, L., Waterfall, H., & Edelman S. (2008). Learn locally, act globally: Learning language with variation set cues. *Cognition*, 109, 423-430.

Owren M.J., Dieter J.A., Seyfarth R.M. and Cheney D.L. (1993). Vocalizations of rhesus (*Macaca mulatta*) and Japanese (*M. fuscata*) macaques cross-fostered between species show evidence of only limited modification. *Dev. Psychobiol*., 26:389-406.

Pelucchi, B., Hay, J. F., and Saffran, J. R. (2009) Statistical leaning in a natural language by 8-month-old infants. *Child development*, 80(3), 674-685.

Puelles, L. (2001). Thoughts on the development, structure and evolution of the mammalian and avian telencephalic pallium. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 356*(1414), 1583-1598.

Puelles, E. Kuwana, E. Puelles, A. Bulfone, K. Shimamura, J. Keleher, Smiga, S., and Rubenstein, J.L.R. (2000) Pallial and subpallial derivatives in the embryonic chick and mouse telencephalon, traced by the expression of the genes Dlx-2, Emx-1, Nkx-2.1, Pax-6, and Tbr-1. *J. Comp. Neurol.* 424, pp. 409–438.

Roper, A. & Zann, R. 2006. The onset of song learning and song tutor selection in fledgling zebra finches. *Ethology* 112, 458-470

Roush, R. S. and Snowdon, C. T. (1994). Ontogeny of food-associated calls in cotton-top tamarins. *Animal Behav.,* 47, 263-273.

Roush, R. S. and Snowdon, C. T. (1999). The effects of social status on food associated calling behavior in captive cotton-top tamarins. *Animal Behav*, 58(6),1299-1305.

Saffran, J. R., Aslin, R. N., and Newport, E. L. (1996) Statistical Learning by 8-Month-Old Infants. *Science* 274 (5294), 1926

Seidenberg, M.S. & Zevin, J.D. (2006) Connectionist Models in Developmental Cognitive Neuroscience: Critical Periods and the Paradox of Success. In Y. Munakata & M. Johnson (Eds.), *Attention & Performance XXI: Processes of Change in Brain and Cognitive Development.* Oxford University Press. pp. 585-612.

Seyfarth, R.M. and Cheney, D. L. (1986). Vocal development in vervet monkeys. *Animal Behav*., 34, 1640–1658.

Seyfarth, R.M. and Cheney, D. L. (1997). Behavioral mechanisms underlying vocal communication in nonhuman primates. *Anim. Learn. Behav.,* 25: 249-267.

Striedter, G. F. (1997). "The telencephalon of tetrapods in evolution." *Brain Behav Evol* 49(4): 179-213.

Tamis-LeMonda, C. S., Adolph, K. E., Lobo, S. A., Karasik, L. B., Dimitroupoulou, K. D., & Ishak, S. (2008). When infants take mothers' advice: 18-month-olds integrate perceptual and social information for guiding motor action. *Developmental Psychology, 44*, 734-746.

West, M. J., and King, A. P. (1988). Female visual displays affect the development of male song in the cowbird. *Nature*, 334(6179):244-246.

Williams, H. (1990) Models for song learning in the zebra finch: fathers or others?, *Animal Behaviour,* Volume 39, Issue 4, April 1990, Pages 745-757

Winter, P., Handley, P., Ploog, D., and Schott, D. (1973). Ontogeny of squirrel monkey calls under normal conditions and under acoustic isolation. *Behaviour,* 47, 230-239.

Wise, R.A. (2006). Role of brain dopamine in food reward and reinforcement. *Philos*. *Trans. R. Soc. Lond. B Biol. Sci*. 361, 1149–1158.

CHAPTER 2

THINKING OUTSIDE THE CORTEX: SOCIAL MOTIVATION IN THE EVOLUTION AND DEVELOPMENT OF LANGUAGE

Abstract

Alteration of the organization of social and motivational neuroanatomical circuitry must have been an essential step in the evolution of human language. Development of vocal communication across species, particularly bird song, and new research on the neural organization and evolution of social and motivational circuitry, together suggest that human language is the result of an obligatory link of a powerful cortico-striatal learning system, and subcortical socio-motivational circuitry.

Introduction

The essential modification in human evolution enabling language has been addressed for decades. In adults, language is demonstrably dependent on the cortex, so the essential evolutionary modification for language has reasonably been sought in the quality or quantity of cortical computations found there. The crystallized product of a lifelong learning process may not be the best source of insight, however, about what initial alterations permitted it. Language plasticity after early damage raises the most perplexing questions about what in the cortex allows language. The right half of the cortex may be removed, or the left half, and language still develops. The frontal cortex may be removed or severely reduced. Uncomfortably for capacity

arguments, total cortical volume may be reduced wholesale by prenatal genetic accident, such as Down's syndrome, or by postnatal accident, and yet language will often survive (Bates et al., 2001; Bates, 2004). An individual may be deaf, blind, or blind and deaf together, and with dedicated tutelage, will still learn language with what remains. What conceivable sensory specialization or cortical alteration could survive these assaults?

We will remain agnostic on the cognitive and cortical architecture of language. Instead, we will highlight the changes in human sociality and its motivational correlates, and explore motivation's mechanistic links to language. Of course, unique aspects of human social structure have already been explored widely and at multiple levels. A general coupling between growth of the neocortex and increasing social complexity in primates has been observed (Dunbar & Schultz, 2007). Demonstrations of how human cultures have solved the problem of stabilization of non-kin altruism, that is, how to stabilize social structures and eliminate cheating such that the manifest benefits of extended altruism to social groups can be had, have been exploded in number (Wilson, Van Vugt, & O'Gorman, 2008; Bowles, 2006; Richerson & Boyd, 2005; Warneken & Tomasello, 2008, Pagel, 2009). In development, human children, contrasted with young chimpanzees, attend to social cues, share information, join games and generally cooperate, serving a form of social learning seen only in humans (Moll & Tomasello, 2007; see also Preston & de Waal, 2001). What is new, and what we are now able to do is place human social behavior more accurately in a comparative and neuroanatomical context, as a transformation of our understanding of the neural circuitry of sociality and its evolutionary variation has occurred (Newman, 1999; Goodson, 2005) (Figure 1).

Making explicit this paper's goals: why evolution and development of social communication in non-human animals should directly inform language development and evolution in humans We will explore the recent advances in the understanding of the social behavior network and its evolutionary variations, highlighting some key findings for those interested in language learning and comparative cognition. We will make some specific proposals about what kind of neuroanatomical circuitry might be changed to link communication with the most fundamental pleasure and reward in human infants. Further, we will hypothesize that such a link constitutes a socio-motivational "gate" necessary and sufficient for the evolution of song in a small brain, and perhaps language in a very large one.

The functional neuroanatomy of bird song will play a central role in this argument. While avian vocal nuclei have often been were first equated with striatal (Edinger, 1988-1903) and then cortical (Jarvis et al., 2005) mammalian domains, we highlight recent evidence that indicates that much of the avian forebrain is in fact limbic (Medina, 2007; Box 1; Figure 2) in nature and is thus better equated with mammalian brain regions associated with social motivation. We argue that this better placement of these structures highlights the intrinsic coupling of sight, hearing and vocalization with motivation highlights the possibility of similar components, largely unexplored, in the neural circuitry of human language learning (Figure 3).

We will be discussing far more empirical detail in evolutionary biology and comparative neuroanatomy than is typical for a developmental science audience. In order to win the reader's patience for this enterprise we argue that new work in development and evolution, "evo-devo", has changed our understanding of the relationship between the behaviors of existing animals profoundly (Kirschner & Gerhart, 2005). What evolutionary presuppositions do we bring to understanding roughly "homologous" behaviors in extant animals, in this case, song in birds

versus language in humans? Contrast two scenarios of behavioral evolution, which have at their base two different theories of brain structure and development. In the first, which we will call the "strict adaptation" model, we imagine an evolutionary random walk from the stem, lizard-like animal to a songbird on the one hand, and a speaking human being on the other. In this scenario, as every feature of the animal's structure and function can be modified in the evolutionary process, the best characterization of any nervous system is of a collection of special-purpose devices (Cosmides & Tooby, 1992) progressively adapted and exapted over the course of evolutionary time. By the time we compare the nervous systems of the bird and the human, such similarities as we would see would indicate only some distant biases caused by the initially shared nervous system plus the functional similarities the similar problems of producing and decoding a primarily vocal signal might produce. The functional organization of bird song might be interesting, but not particularly directly instructive about human language learning.

On the other hand, consider a different scenario we will call "developmental conservation", which has arisen from the consistent observation of startling similarities in the basic body and brain plans across vertebrates (Kirschner and Gerhart, 2005). In this scenario, before divergence of birds and mammals, a basic, robust, and evolvable brain plan had stabilized. Certain kinds of variations, those associated with useful adaptations and niche exploration in the past, were facilitated, and embedded in epigenetic programs that produced stable function in commonly encountered environments. In this scenario, only certain locations in the nervous system can produce vocal behavior, only certain others can link vocal behavior to particular social settings, only certain specific regions can be the site of long-term associative learning. In the "developmental conservation" model, the functional similarities of song and language are amplified by a conserved basic brain plan and epigenetic operations stabilized over time to

facilitate only certain kinds of operations and adaptations. Much debate remains about just how much evolution resembles the first versus the second scenario, and we do not pretend the issue is settled! Here, however, we are making the strongest possible case for "developmental conservation". Not only because we are convinced by the increasing evidence for it, but also to model this relatively new argument for the developmental community who are neither evolutionary biologists nor anatomists.

In the following, we will review evidence for motivational dependencies in learning of social communication. We will also review briefly the neuroanatomy thought to be important in essential motivation, including what has been discovered recently in evolutionary variation in motivational circuits. We will discuss some mechanisms for recognition and learning from caregivers and argue that small biases in early attention and motivation support the multimodal initial recognition of and eventual complex representation of conspecifics and caregivers. We will conclude with specific suggestions about where to look in the developing brain for the structural and physiological changes which seem likely. In sum, we attempt to return attention to language in its ecological context, acquired in the early development of an obligatorily social, gregarious, and often-altruistic species.

Situated, Social, Motivated Vocal Behaviour

Cheng and Durand [2004] have defined emotion as an "all inclusive state, mediated by endocrine, autonomic, and skeletomotor responses organized by pathways through the amygdala and the hypothalamus" and have argued that "if vocal behavior is recognized as an interdependent skeletomotor component of emotional state, then its dependence upon internal stimuli (neuroendocrine and autonomic activity) for expression can be appreciated; singing is a

motivated behavior".

Similarly, vertebrates who (independent of learning) vocalize to define territories, or announce their species identity and distance, or readiness to mate, or express pain or alarm obviously do so in an explicit motivational state. Interestingly, even in organisms that don't need to *learn* their species-typical vocalization (Owren Dieter, Seyfarth & Cheney, 1992; Seyfarth & Cheney, 1986; Winter, Handley, Ploog, & Schott, 1973), social context typically modifies and modulates the nature and the content of such vocalizations (Roush & Snowdon, 1994, 1999; Seyfarth & Cheney, 1997). The vocal learning trajectory of species that do not explicitly require social interaction in order to produce species typical vocal forms could also be seen to provide a way in which claims of the centrality of social motivation in vocal learning can be falsified. For instance, it is known that territorial songbirds, such as the white-crowned sparrow, *Zonotrichia leucophrys*, are able to reproduce species typical adult song on the basis of exposure to taped tutor song during the sensitive phase for song learning (Marler, 1970) i.e. in the absence of social interaction. However, immense power that social motivation exerts on learning is demonstrated in the work of Baptista and Petrinovich (1984), who showed that a young white-crowned sparrow will learn additional song elements from a either a conspecific or a heterospecific live tutor even after the song learning period. Similarly, numerous additional examples underscore the potency that social stimuli exert on vocal behavior. For instance, changes in social structure will cause adults birds to modify their non-song vocalizations in a phenomenon known as *vocal sharing* whereby individuals change their own vocalizations to resemble those of their conspecifics. Such is the case with black-capped chickadees (*Poecile atricapillus*) that modify the structure of their call at the time of formation of the winter flock, with all birds converging towards a common structure (Nowicki, 1989). Similarly, when the social groups of starlings

(*Sturnus vulgaris*) are changed, they change their song type to match that of their companions (Hausberger, Richard-Yris, Henry, Lepage, & Schmidt, 1995). Vocal sharing is used not only to initiate and maintain social bonds within units, but also to signal alliance to compete against other units, as in bottlenose dolphins (*Tursiops truncatus*) (Smolker & Pepper, 1999).

Vocal sharing in a changed social context has been seen in various non-human primates as well. (Snowdon & Elowson, 1994) found that within 10 weeks of being housed together, all the animals from two different groups of pygmy marmosets changed their call structure in the same direction, regardless of reproductive status. Snowdon et al. (1997) also found that mate pairs of pygmy marmosets (*Cebuella pygmaea),* who had stable trill-structure for months before pairing, significantly changed their trill structure in the 6 weeks immediately after pairing and then this changed structure was stable for at least 3 years after pairing. Female squirrel monkeys produce vocalizations known as "chucks" with individualistic acoustic properties that constitute a "vocal signature," and are used to preferentially respond to familiar group members, but not to acoustically similar chucks of unfamiliar conspecifics (Smith, Newman, Hoffman, & Fetterley, 1982a; Biben & Symmes, 1991). These examples illustrate the intrinsic role that the nature of the social context plays in animal vocalizations.

More elaborate vocal learning, defined as "the ability to modify acoustic and/or syntactic structure of sounds produced, including imitation and improvisation"(Jarvis, 2007), is a phylogenetically more limited phenomenon. Vocal learning has been unequivocally attributed to 7 taxa in all, of which 3 are birds - songbirds, hummingbirds and parrots, and 4 are mammals - cetaceans, elephants, bats and humans (Jarvis, 2007). In organisms that learn a complex vocal communication system, social context is typically necessary for this process. Deprivation studies were the starting point for this research. Studies of human children (Fromkin, Krashen, Curtis,

Rigler, & Rigler, 1974; Lane, 1976) or songbirds (Marler, 1970; Thorpe, 1958) that were raised in isolation showed that vocal learning from conspecifics is necessary for the development of normal speech. It is noteworthy that both the rudimentary speech of children and the isolate song of birds reared in non-social setting, while simplified, do show some species-specific features (Marler & Sherman, 1985). While this is most often cited as evidence for either an innate model for language (Chomsky, 1981; Fodor, 1983) or an innate human capacity for learning to segment and group sensory stimuli (Elman et al., 1996), there is an interesting way to recast this finding. Perhaps the underdeveloped vocal repertoires that characterize these infants and nestlings show what is obtained when motivation is subtracted from the vocal learning equation, and thus seriously misrepresent the structure and preferences which both bring to vocal communication.

.

The Social Circuit

Until the last ten years, the circuitry of the brain central to social behaviour was described as a collection of subcortical centers and areas with specific labels, i.e. male aggression, female parental care and so forth. Integrating across an enormous number of studies, Sarah Newman (1999) proposed a new formulation of the "social network" in the mammalian brain comprised of six heavily interconnected limbic system areas with interlaced functions (Figure 1).

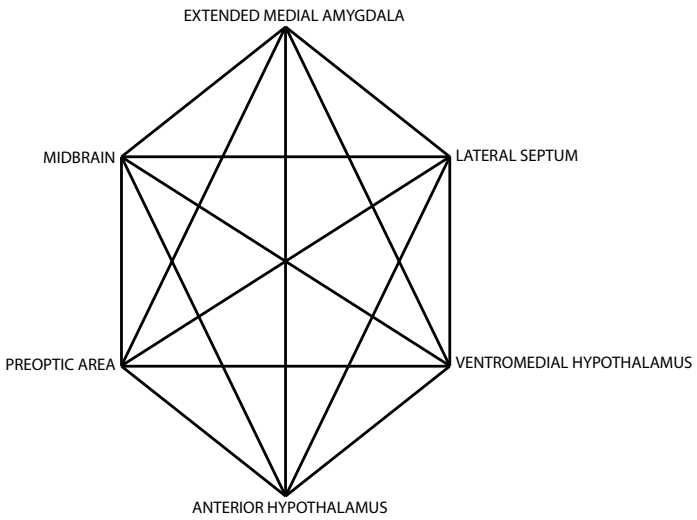

Figure 2.1 The social behaviour network as originally described by Newman, 1999

In this conception, any given social behaviour is an emergent property of the temporal and dynamic pattern of activity across this social network in the brain, replacing the notion that each node is a centre for some category of motivated behaviour. Rather, varying levels of activity across the nodes signify a given behaviour. The sensory stimuli important for these behaviours differ with species, sex, maturational state and context, as do the motor manifestations vary with each animal's competence. Ecological and social niches, such as gregarious versus territorial or diurnal versus nocturnal systems vary as well, and the rather astonishing observation is the phylogenetic stability of the location of this system in the face of such sensory, motor and niche variability (Newman, 1999 and references therein). Also stable, are the essential functions of the social circuit: mate choice and mating, parental behaviour, territoriality and aggression, affiliation and so forth, while the instantiation of each function in a particular body and ecology varies. This polysensory, distributed nature of the social network may offer a way over evolutionary time to modulate sensory preferences and dispositions with relative ease while conserving core functionality.

Using this framework, a discussion can begin of how such a neural system evolves to produce varied vertebrate social behaviours, how vital functions are paired with the sensory systems each animal best employs, its motor abilities and motivations. There is already evidence to indicate that the dynamic activation of this network can "code" for the (negative or positive) valence associated with social encounters. Goodson et al. (2005) located a set of closely-related finches which varied in their preferred social context, from solitary and territorial to flocking, but had similar mating systems. They found species-specific differences in immediate early gene (IEG) expression in the medial extended amygdala following exposure to a same sex conspecific. Activity in this region was negatively correlated with species typical group size – greater activation of the extended amygdala was associated with the response of a territorial species to same-sex conspecifics – a negatively-valenced social encounter, while there was less activity if the bird welcomed the "intruder", as in the case of the gregarious species.

Both behavioural differences and brain activity differences can also be related to neuropeptide receptor distributions. While the brain distributions of the actual neuropeptides such as oxytocin and vasopressin are very highly conserved across species (Wang, Zhou, Hulihan, & Insel, 1996) the location and density of the receptors for these peptides is not conserved (Young, Nilsen, Waymire, MacGregor, & Insel, 1999). The evolutionary alteration of receptor densities has been well described in the case of two closely-related voles. Some voles are socially monogamous, sharing parental care; in these voles, copulation with a partner fixes an enduring preference for the partner's company, not shared by closely related promiscuous species. In the monogamous species, expression of a receptor for vasopressin in reward areas of the basal forebrain permits a link to be forged between the specific representation of an individual vole signaled by vasopressin activity, and reward and comfort itself. With the

presence of the vasopressin receptor, a gate is opened between the sensory representation of the individual mate and basic reward, stronger than the general links that can be made between any random cue and a positive outcome.

Specifically, in the monogamous prairie vole, vasopressin receptor (V1aR) binding (concentration) is very dense in the ventral forebrain, in particular in the ventral pallidum which is a region associated with reward and motivation (Lim, Murphy, & Young, 2004). Injection of vasopressin into the brain, paired with presentation of a female vole increases the tendency to affiliate with that female, while blocking vasopressin activity abolishes established partner preference (Lim & Young, 2002, 2004). On the other hand, non-monogamous montane and meadow voles naturally have very low densities of V1aR receptors in this region (Insel, Wang, & Ferris, 1994) but experimentally increasing the number of these receptors can lead to pair bond formation even in these promiscuous species (Lim, Hammock, & Young, 2004).

Differences in neuropeptide activity within the social network also underlie other kinds of social preferences in birds. Goodson & Evans (2004) studied the social network in a range of species of birds that differed in their coloniality. Different levels of colocalization of AVT (Arginine VasoTocin-the avian orthologue of mammalian vasopressin) and immediate early gene expression in the medial bed nucleus of the stria terminalis (BSTm) and lateral septum marked the varying evaluation of particular social situations by bird species with different levels of gregariousness. The design of these experiments did not permit overt social behaviour, thus highlighting evaluation and motivation rather than action. In essence these data indicate that presenting a bird with stimuli that are positively valenced within its ecological niche causes an increase in uptake of AVT, and immediate early gene activity, whereas the opposite is true for negatively valenced stimuli. So for instance, to a gregarious bird, both partners and same-sex

conspecifics present opportunities for rewarding social interaction and a corresponding increase in AVTir-IEG colocalization is seen in these birds upon presentation of either of those classes of stimuli. To a territorial bird on the other hand, a same sex conspecific signals a potential antagonistic encounter – and this is associated with a decrease in AVTir-IEG colocalization whereas the presence of the subject birds partner – presumably a positive stimulus – causes an increase in colocalization.

Whether the effects of vasopressin occur directly through association with reward pathways in the ventral pallidum or through the neuroendocrine mechanisms associated with stress and social anxiety (Goodson & Evans, 2004), the activity of this neuropeptide seems to provide a robust system under which the evolution of sociality can be probed. In effect, one can think of the receptor distribution and concentration of vasopressin receptors as locks that can migrate and vary in number, with keys always standing ready. In experiments, changing the level of only one receptor appears able to produce a change in social preference enough to alter a species' life history and species-level social structure. Vasopressin receptor levels alone, of course, probably are not the only distinctions between social species from non-social species over an evolutionary time-scale. The relative simplicity of the structural changes necessary to bring about large scale changes in the ecological niche of a given organism are, however, impressive, and candidate genes have already been described, in the case of the voles (Young, Nilsen, Waymire, MacGregor, & Insel, 1999; Hammock & Young, 2004, 2005)

Obscure objects of desire: representing an individual

Before proceeding, it is worthwhile to stand back and consider some comparative information on what "specific representation of an individual" may entail. Early findings in

ethology characterized interactions between animals as specially evolved stimulus-response linkages, such as herring gulls pecking at the red spot on the feeder's beak (Alcock, 2009). These early examples of stimulus-response singularities do not characterize the representational complexity of most vertebrate social interactions particularly well. The research that arose from these first examples showed that a large range of mammals (and probably, most vertebrates) come to be endowed with a far more robust representation of a conspecific than a single sensory emblem. The corpus of work by Eric Keverne and colleagues on maternal imprinting on lambs is instructive (Broad, Curley, & Keverne 2006). Flocks of sheep give birth relatively simultaneously in the spring, and the ewe has a specialized, time-limited postnatal learning period, hormone dependent, in which she will learn to recognize her own lamb. Allowing the lamb at first to suckle is dependent on olfaction, but the ewe must solve the greater problem of recognizing her own lamb amongst others within a day or so. While olfaction can be involved in this, it is not essential, and all relevant modalities are brought to bear, the bleat, the face and body appearance. The object of maternal desire is not a value in single sensory channel, but an individual lamb. Similarly in rodents, the various separable odorants associated with an individual, encountered in different contexts, serve as tokens of the whole (Johnston, 2005). The social behavior network may be seen as the conserved neural structure that assembles on the one hand relevant sensory dimensions into a representation of an individual (either by innate bias or experience-expectant learning), and on the other attaches that representation to the motivations and actions appropriate to their social roles, parent, child, mate or rival. With situation, and maturation, these attachments also change as required.

Turning back to language evolution, we can use this framework to better understand how, even in closely-related species, the motivational significance of a particular conspecific role

could differ radically, depending on which gates are opened to reward and comfort, or to anxiety and aversion. These simple, but significant changes can immediately alter the entire social setting and preferences of an individual, who they attend to and stay close to. The increased social orientation of human infants to solicit other humans' play and help, compared to chimpanzees is exactly the type of change we are attempting to provide a mechanistic account for (Tomasello, 1999). The utility of such changes to motivate language learning seems reasonable, but social circuitry *per se* as we described it so far carries no explicit way to potentiate particular learned behaviours.

Links between the social circuit and motivation systems

It is worthwhile taking a moment to discuss "motivation" and it's relationship to processes of social cognition and reward. Historically, Behaviorists have studied motivation in the context of stimulus-reward contingencies whereas Social Psychologists have studied the cognitive processes that underlie social behavior and reward. It is no surprise then, that role of motivation has found much more favor in non-human animal communicative behavior, which is considered more instinctual and less cognitive, and has formed an integral part of the attempt to understand the processes that govern such behavior. Part of this disparity stems from the piecemeal leftovers of the doctrine of encephalization [Steklis and Raleigh, 1979]. As originally described by Hughling Jackson in 1887, the main tenet of this doctrine was that structural and functional evolution of the vertebrate nervous system proceeds from a a caudal to rostral direction [Jackson, 1888]. Accordingly, phylogenetically more recent or "advanced" species are more reliant on neocortical (rostral) mechanisms for the execution of functions that are associated with more primitive subcortical or limbic structures in the brains of phylogenetically

more ancient or "lesser" species. Thus the common thought is that while motivations (encoded in subcortical and limbic domains) form a central part of vocal behavior that is largely associated with stereotypic and sometimes involuntary communication of affective states in other animals (such as non-human primates) [Rowell and Hinde, 1962], the more highly developed "voluntary" language mechanisms of humans must lie in the more recent and purportedly more advance neocortex. The purpose of this paper is not to undermine the advancement of the cortex or its functions in language associated processing - but an important point to consider here is that a linkage of linguistic function to motivational or emotional states does not preclude it being subject to either higher order voluntary control or being encoded (in part) in neocortical domains.

Recent studies have indicated that physiological properties such as opioid neurotransmission of neural circuits that underlie motivation sub-serve social reward (*Trezza* et al., 2011) and also that traditional motivation regions such as the striatum are activated in response to both non-social and social reward (Izuma et al., 2008). Thus it appears that social reward is the affective implication of social motives wherein affective states can serve as both a source of motivation for behavior (Forgas and Laham, 2005) as well as provide feedback about the status of motivated behaviors (Carver and Scheier, 1998). Within the social system of motivation then, sociality becomes the motivating force behind executing a behavior – in this case, vocal behavior, or language.

A mountain of evidence exists in support of the claim that limbic structures (and the hypothalamus) code for motivational processes and the reader is referred to many excellent reviews for further reading (Mogenson, Jones, & Yim, 1980; Cardinal, Parkinson, Hall & Everitt, 2002). For the purposes of this essay, we wish to highlight the overlap between much of the social circuit and parts of the motivation/emotion circuits (Figures 2 and 3). In the large

expanse of the limbic forebrain, the extended amygdala forms a relay center for information crossover between the cortical and brainstem regions. While it acts in concert with other regions of the brain in encoding emotional processes, the amygdala, via its modulation of both attentional and motivational arousal, is one very salient functional seat of motivation in the mammalian brain. The connections between the amygdala and the ventral tegmental area, via the nucleus accumbens - a structure located in the ventral striatum - form a large part of the mesolimbic dopamine pathway that modulates behavioral responses to stimuli that activate feelings of reward or motivation. According to the Depue and Strupinsky (2005) model of affiliative bonding in humans, the convergence of dopaminergic, neuropeptide and opiate systems in the ventral striatum provides for the human capacity to feel rewarded by affiliative (social) stimuli.

While additional regions of the mammalian brain have been associated with motivation-related processes (hippocampal formation, the cingulate cortex, and limbic midbrain areas (Morgane, Galler, & Mokler, 2005) one that bears mention here is the prefrontal cortex (PFC). The pre-frontal cortex regulates complex memory, attention and cognitive processes and is also an integral part of the motivation circuit via its extensive bi-directional connectivity with the limbic forebrain (Damasio, Grabowski, Frank, Galaburda, & Damasio, 1994; Goldman-Rakic, 1999). In humans, increased gray matter density in the orbitofrontal cortex (PFC) and the ventral striatum (Lebreton, et al., 2009) and increased connectivity between the frontal cortex (including the orbitofrontal cortex) and striatum (Cohen, Schiene-Bake, Elger, & Weber, 2009) are associated with increased responsiveness to socially defined reward and increased disposition to social relationships and attachments.

So how do these structures tie into vocal learning and/ or production? It is instructive to

think of vocal learning as a specialized learning behavior that, like most (if not all) behavioral systems that incorporates any kind of learning, exists in a motivational context. The intimate linkages between the motivation and social circuits provide us with a system that can represent and attach motivational value to social conspecifics. All that is required, over evolutionary time, is the linkage of vocal centers to that social-motivation system - and together these will comprise the vocal learning system. (Figure 3A and 3B) And these links are found at both the anatomical level and at the behavioral level. For instance, projections of the ventral tegmental area to the avian vocal centers modulate early gene activity related to directed or undirected singing in zebra finches (Hara, Kubikova, Hessler, & Jarvis, 2007) The extensive work on birdsong, and the fairly recent re-understanding of the homologies between the neuroanatomical regions involved in vocal communication in birds and in humans will be our vehicle to a mechanistic understanding of where brain changes may have occurred.

Birdsong and human language

Many parallels between birdsong and human speech have been explored; perhaps the most striking is the interplay between production and perception in the development and maintenance of vocal communication (Doupe & Kuhl, 1999). The parallels in comparative neuroanatomy, critical periods, mechanisms of learning, lateralization and so on are also striking. From this voluminous literature, we will concentrate on those studies that emphasize the natural ecology of learning, its neuroanatomical motivational substrates and their interaction.

Although a full discussions is beyond the scope of this review, the renaming of the structures of the bird telencephalon according to their proper site of origin, neuropharmacology and connectivity to allow proper homologies to other vertebrate telencephalon (including our

own) has brought many "bird song nuclei" into the basal forebrain motivational circuitry where formerly they had been given quasi-cortical status. Under this scheme (Medina, 2007; Box 1; Figure 2), a number of key nuclei in the song control system now fall in the region of brain homologous to basal forebrain and amygdala. In Figure 2, a-d, the divisions of the embryonic telencephalon which give rise to the hippocampus, dorsal pallium and isocortex; nidopallium, entorhinal cortex and amygdala; arcopallium and the striatum, as well as the thalamic-hypothalamic core are given. Looking at the location of birdsong nuclei in this light, it is clear that most fall into those regions associated with direct access to motivational and social circuitry, and not the neocortex. By the argument we made earlier about the essential conservation of fundamental operations in the vertebrate brain, if much of the circuitry of this learned, social, vocal behavior falls in the extended amygdala and associated limbic regions in birds, we should seek its homologue(s) in the other vertebrate vocal learner we commonly study, the human infant.

***BOX 2.1: A note on developmental neuroanatomy*:** The vertebrate forebrain is divided into the telencephalon and the diencephalon. The diencephalon contains the regions described by the thalamus and the hypothalamus. The telencephalon is divided into a dorsal pallium and a ventral subpallium. The subpallium develops into the basal ganglia and related structures whereas the pallium forms the cortex and parts of the amygdala. The pallium can be further divided into medial, dorsal, lateral and ventral pallium (Puelles et al., 2000). In mammals, the lateral and ventral pallia (originally together referred to as lateral pallium) form the piriform cortex, claustrum and parts of the amygdala. Much of the limbic/motivational circuitry in mammals derives from the lateral and ventral pallia with the medial pallium giving rise to the

hippocampal formation. The mammalian isocortex derives from the dorsal pallium. Cross-species comparisons of these subdivisons based on only adult anatomical and functional data have led to the misplaced notion that the latero-ventral pallial regions of the avian brain are homologous to parts of the mammalian isocortex (Reiner 1993; Karten, 1997), when the only region of the adult birds brain that the dorsal pallium gives rise to seems to be the "Wulst" (Medina, 2007). Data derived from embryological origin suggests that most of the avian forebrain comes the latero-ventral pallium (also called the DVR or the dorsal ventricular ridge) which is homologous to the mammalian claustrum and pallial amygdala (Striedter, 1997; Puelles et al., 2000; Puelles 2001; Martinez-Garcia, Martinez-Marcos, & Lanuza, 2002; Martinez-Garcia, Novejarque, & Lanuza, 2007). This would indicate that the nidopallial vocal nuclei of avian species are more soundly positioned in regions homologous to the limbic forebrain of mammals – rather than regions akin to the mammalian neocortex.

The reason this is important is that location often determines function via the class of connections a region makes, the neurotransmitters it expresses etc. Mislocalizing" a lateral pallial derivative (a location which can project directly to hypothalamus and forebrain, "motivational" circuitry) as a dorsal pallial structure (which does not have this class of connections) could cause us to misconstrue its functional roles.

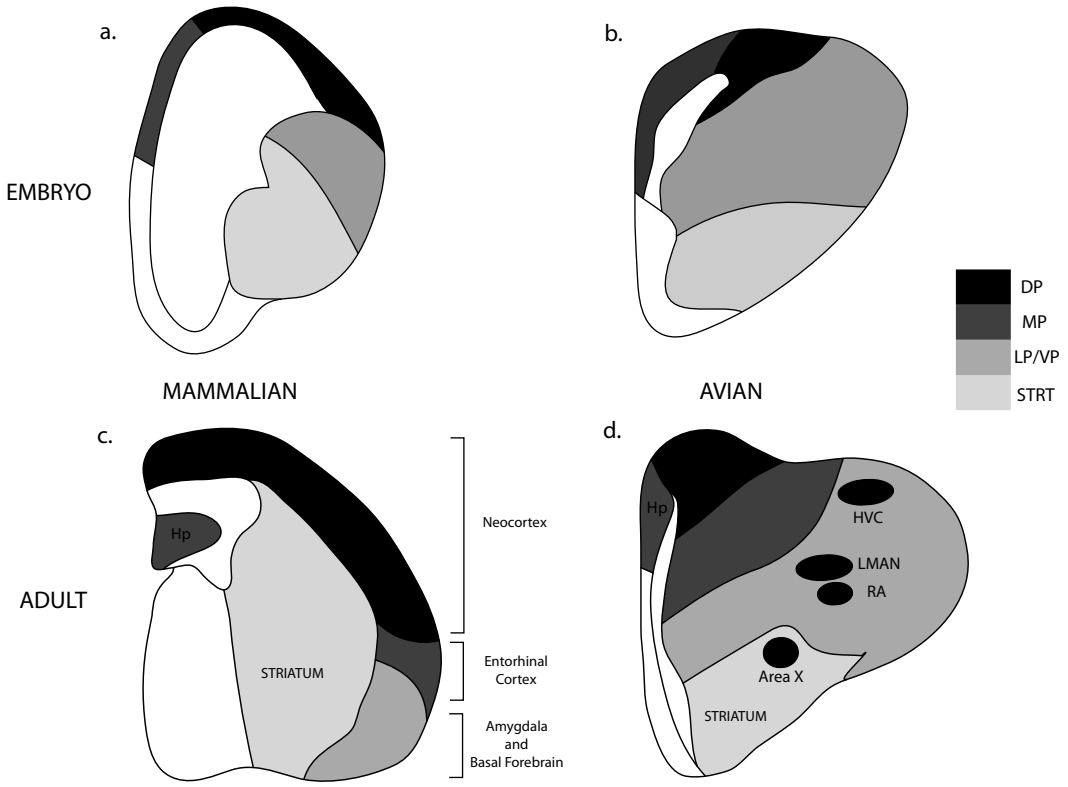

Figure 2.2 Schematic of transverse sections through mammalian and avian brains at two stages of development (Modified from Fig.2, Medina, 2007)

2a) Mammalian embryo, 2b) Avian embryo, 2c) Mammalian adult, 2d) Avian adult with the major vocal nuclei delineated according to embryonic tissue origin. Note that multiple sections along the anterior-posterior axis have been collapsed in order to allow visualization of all vocal nuclei within the same section. Abbreviations: DP – Dorsal Pallium, Hp – Hippocampus, MP – Middle Pallium, LP – Lateral Pallium, VP – Ventral Pallium, STRT – Striatum. Vocal Nuclei: HVC – High Vocal Center, LMAN – Lateral Magnocellular Nucleus of the Nidopallium, RA – Robust Nucleus of the Arcopallium. Only the dorsal pallium gives rise to neo-cortical tissue in mammals (shaded in black). The bird vocal nuclei are not described by this region but instead all lie within tissue derived from lateral and ventral pallia (along with the striatum) which give rise to much of the motivational/social circuitry i.e. the amygdala and the basal forebrain in mammals. Note that the prevailing view in the field of bird song still holds that pallial song nuclei are transcribed by mammalian cortical equivalents. We wish to turn attention to the growing body of embryological and topological data that provides emphatic evidence that song nuclei are limbic/motivational in origin (Box 1).

In 1969, Immelman conducted what is now held as a seminal experiment in the progression of the science of birdsong learning. He cross-fostered infant zebra finches in Bengalese finch families and found that the zebra finches learned the song of their heterospecific fathers even when adult conspecific males could be heard nearby (Immelman, 1969). Since then, zebra finches have been also shown to need social interaction with a real-bird tutor (most often, the father) in order to learn a song from him and typically don't learn well from tapes (Eales, 1989; Mann & Slater, 1995). If, however, the experimental setting is such that a subject bird is required to activate a tape by pressing a key to elicit the zebra finch song model, song is learnt. Yoked controls who can hear the same playback, however, do not learn the song (Adret, 1993). The commonality between a live tutor and the controllable playback is response contingency. While the results of this particular study are contentious because a subsequent study was not able to replicate these effects specific in zebra finches (Houx and ten Cate, 1999), there is now ample evidence from human studies to indicate that response contingency is crucial to infant learning of advanced vocal forms (Bloom, Russel & Wessenberg, 1987; Goldstein & Schwade, 2008). In addition, response contingency in the form of turn taking interactions is an early feature of parent-infant face-to-face interactions (Papousek, Papousek, & Bornstein, 1985) and human infants respond to social contingency from a very young age and show sensitivity to familiar contingency levels based on the responsiveness of their caregivers (Bigelow, 1998; Bigelow & Rochat, 2006). It has been suggested that social contingency aids learning by focussing infant attention on relevant features of stimulus information (Kulh, 2007). It is thus a property of the interaction itself, and not sensory signals or motor output that gates learning, an observation that is in fact fairly pervasive. Similarly, parrots who are trained to produce a variety of words and labels in human language, do not learn through watching videotapes or listening to audio tapes or

by watching non-interactive humans (Pepperberg, 1999).

In zebra finches, neural activity during singing is strongly modulated by the presence of a listener. Both the magnitude and the variability of activity in the LMAN (lateral magnocellular nucleus of the nidopallium, song nucleus) and Area X (basal ganglia, song nucleus) is dependent on the presence of a conspecific - the activity is lower and more uniform during directed singing (to a female) and higher and more variable during undirected singing (alone) (Hessler & Doupe, 1999). Dopamine levels in the Area X are significantly higher during directed singing as compared to undirected singing - possibly because of higher dopamine release from the VTA-SNc (Ventral tegmental area – Substantia nigra pars caudalis – reward region) terminal that project to Area X (Sasaki, Sotnikova, Gainetdinov & Jarvis, 2006). Area X is necessary for song learning and is part of the AFP (anterior forebrain pathway pathway: cortex → basal ganglia → thalamus →cortex), which is involved in auditory feedback based evaluation and adaptive modification of the bird's own song. High dopamine levels in the striatum of birds during directed song parallel high dopamine levels in the striatum after drug administration in mammals (Sasaki, Sotnikova, Gainetdinov, & Jarvis, 2006) – pointing to the rewarding nature of the "feeling of being heard". Increased glutamate activity in the VTA is also seen after intense social interaction (Huang & Hessler, 2008). The significance of social context in vocal behavior is further highlighted by the observation that new neuronal recruitment to the HVC (high vocal center, song nulceus), Area X and NC (nidopallium caudale, cortex) increases in the brains of adult males who are placed in a complex social context (large aviary setting) as compared to males who are placed in a simple social context (pair cages) (Lipkind, Nottebohm, Rado, & Barnea, 2002). This is true even though birds in the simple social environment sing much more than the birds in the complex environment, indicating that the neuronal recruitment is associated

with the processing and storage of rich auditory input in a rich social setting rather than mere vocal exertion (Adar, Lotem, & Barnea, 2008).

The concept of intrinsic reward in vocalization (Cheng & Durand 2004), with direct physiological outcomes such as ovulation, has been studied extensively in ring doves (*Streptopelia risoria)*. In the female ring dove, Estrogen acts on a midbrain song nucleus (mICo) to induce a female nest coo in response to a male courtship coo (Cohen & Cheng, 1981). Both auditory and proprioceptive feedback from the female bird's own cooing (and not the male's cooing) initiates an endocrine cascade that ultimately results in egg-laying (Cheng, 2003). In brown-headed cowbirds (*Molothrus ater*), the female provides feedback to a singing male via a wingstroke when she likes his song - something that causes him to incorporate more of the "approved" acoustic forms into his vocal repertoire (West & King, 1988). Such a gesture on the part of the female cowbird could act like the cooing vocalization in ring doves, whereby the wing stroke, acting via the circuit that receives modulatory affective input from the arcopallium/amygdala, could not only be a signal of the willingness of the bird to copulate, but in fact self-stimulate her to do so.

In summary, the extensive and expansive connectivity that exists between motivation systems and the vocal nuclei in avian species, along with behavioural data that implicates affective systems in vocal learning processes, suggests that neural pathways that mediate emotional and motivational states and vocal control pathways form an *integrated* system of communication in birds. Further, these pathways receive constant direction from "social road-signs" within the birds' ecological niche. In the avian species that display learned vocal behaviour, the learning environment is best viewed as a consolidated system that incorporates the entire infant bird, its tutor, the interaction between the two, and the effect of each actor on its

own, and the other's nervous system.

Social motivation in human language learning

The role of motivation or affect during communicative interactions between human mothers and infants has long been appreciated. Factors such as the prosody of caregivers' early speech to infants, along with facial movements, form non-linguistic affective behaviours that "attract and sustain the infants' attention to what will become a linguistically important stream of cues" (Locke, 1993). Caregivers use infant directed speech to facilitate infant arousal and motivation, which in turn facilitates infant learning of language. The prosody of infant directed [ID] speech is different from adult directed speech. Caregivers' speech to infants is often characterized by shorter utterances, longer pauses, higher pitch and wider pitch excursions (Fernald & Simon 1984; Fernald et al., 1989). Infants can discriminate affective vocal expressions in ID speech in their own language as well as foreign languages and the intonation of ID speech can elicit emotional responses from infants (Fernald, 1992 a, b). The intonation of ID speech has been shown to effectively engage infant interest (Kuhl, Coffet-Corina, Padden, & Dawson, 2005). Finally, mothers reliably use falling pitch contours in speech to infants when trying to soothe (Fernald, Kermanschachi & Lees, 1984) and rising pitch contours to when trying to engage infant attention (in a social interaction) (Stern, Spieker, & MacKain, 1982). These data have led to the idea that maternal prosody is finely tuned to infant attention and arousal level (Fernald, 1985) and can be used to modulate infant motivational states. In further support of the claim that motivation drives acquisition of communicative faculties, prelinguistic vocal learning in human infants has also been shown to be modulated by social reinforcement. Contingent responsiveness from caregivers, both silent and vocal, is associated with vocal learning in human

infants (Goldstein et al., 2003; Goldstein & Schwade, 2008). Further infant learning of phonemic contrasts in a foreign language is contingent on their presentation by a socially interactive live tutor (Kuhl, Tsao, & Liu, 2003). Based on this and other work, Kuhl (2007) has advanced the "social gating" hypothesis whereby infant language is gated by the motivating properties (such as attention and arousal) inherent in social interactions.

In some bilingual households, children only acquire proficiency in one language (Wong Fillmore, 1991). It has been argued that for these children, the language that is not learned is the one that is less "significant" to the child, in that it is used less often in settings wherein the child is likely to be motivated, for example among peer or sibling groups.

Aberrant language acquisition

A particularly interesting case of the interplay between motivational and vocal centres in the brain is autism. Whether or not the limbic system dysfunction plays a causal role in the pathology of this disorder remains contentious but there does seem to be some support for the idea that the amygdala and hippocampus of autistic individuals develops abnormally (Schumann et al., 2004; Nacewicz et al., 2006]. Approximately 25% of all children with autism never develop functional language abilities (Klinger, Dawson, & Renner, 2002). Unlike typically developing children, autistic children prefer non-speech samples to acoustically matched samples of ID speech and reportedly display no "interest" in language, which has been correlated with aberrant neural responses to speech. They also have trouble maintaining attention in linguistic tasks (Kuhl, Coffet-Corina, Padden, & Dawson, 2005) and these attention deficits in children with autism are also evident in their inability to form and maintain joint focus of attention (Osterling & Dawson, 1994; Osterling, Dawson, & Munson, 2002). At around 9 month of age,

typically developing human infants start to engage in triadic interactions that involve them, their caregivers and a third external entity that might be a toy or another person or anything in the outside environment. These interactions involve joint attention between the infant and the caregiver and it has been argued that the ability to do so is a necessary milestone in the development of communicative faculties (Tomasello, 1999). In addition to this, overall attention and shared visual attention between infant and tutor predict the degree to which infants will learn the phonemes and words of a language (Conboy, Meltzoff, & Kuhl, 2008). Autistic adults have an exaggerated anxiety response to eyes, and when looking at faces, attend to the mouth, a region less informative about certain emotions, and certainly about where their interlocutor is looking (Neumann, Spezio, Piven & Adolphs, 2006). Further, they have impaired visual memory for faces (Blair, Frith, Smith, Abell, & Cipolotti, 2002) and their neural responses to social stimuli don't show attentional modulation (Bird, Catmur, Silani, Frith, & Frith, 2006). That being said, recent data have indicated that adults with Asperger's syndrome don't make initial saccades or show a looking bias towards the correct choice in a false belief task, even though they don't differ from normals in their performance in a verbally mediated theory of mind task (Senju, Southgate, White, & Frith, 2009). In the same vein, it has been suggested that the correlation between verbal mental age and ability to pass a theory of mind task in children with Autism may be a result of verbally mediated reasoning, rather than social attunement (Happe, 1995). While these data may seem incongruent with the idea of an obligatory link between linguistic ability and proclivity towards social behaviour, it is important to note that these are the results of compensatory mechanisms that come into play during or after wide scale deficits in development. For instance, children who are able to use verbal reasoning to complete a theory of mind task form only a very small proportion of all autistic children and further, often require a

verbal mental age that is double that of normal children who are able to complete similar tasks (Happe, 1995). Thus, the fact that they are able to use some kind of cognitive compensation is indicative of the general functional plasticity of the brain, and is not at odds with the theory that the social motivation that these children seem to lack would have gated earlier (typical) maturation of verbal ability, as well as social reasoning.

Human infants with William's Syndrome have disrupted cortical development, including increased cortical thickness (Thompson et al., 2005) and mild to moderate mental retardation but don't show linguistic disability of a magnitude that corresponds to these cortical anomalies (Bellugi, Lichtenberger, Jones, Lai, & St. George, 2000). As adults, these individuals show heightened sociability (Jones et al., 2000) and interest in faces (Laing, Butterworth, & Ansari. 2002) but don't seem to recruit amygdalar function in a manner akin to normals during face processing tasks (Paul et al., 2009). In general, there seems to a distinct propensity for heightened affect, especially in social situations, such as those associated with linguistic narrative (Reilly, Losh, Bellugi, & Wulfeck, 2004). That there is some abnormality is not being contended, and the argument here is *not* that individuals with Williams syndrome have completely preserved linguistic skills in the face of extensive cortical aberrations that result in retardation, but that given the extensiveness of the cortical damage, the preservation of even the compromised linguistic ability that these individuals display is surprising - and suggests the possible involvement of areas that lie outside the cortex in linguistic processes. Our suggestion is that spared or exaggerated social motivation in these individuals accounts for the *protection* of language.

We would like to note that both Autism spectrum disorders and William's Syndrome are complex and incompletely understood brain disorders and it would be beyond the scope of our

current knowledge to draw definitive conclusions from the data we currently have, either in favor of social motivation as the driving force behind early language learning, or in favor of a disjunct between social and linguistic behavior. However, on the basis of emerging trends from current research, we would argue that studies of individuals with these disorders do seem to indicate that they provide us with populations where either the social system is not the *tour-de-force* that it normally is (Autism) or where hyperactivity of the social circuits is offsetting cognitive deficits in a manner that provides impetus to greater learning of faculties that are embedded within the social domain (William's Syndrome).

Finally, an aside about another genetic disorder associated with language deficiency - the alteration in the expression of FOXP2 producing serious deficiencies in the production of and structure of language in the KE family, as well as multiple deficiencies in praxis. While it is not yet known exactly what feature of brain organization its early FOXP2 expression supervises, the fact that it is under positive selection in both songbirds and humans, is expressed in the basal ganglia, and associated with some aspect of fine temporal organization in both production and perception is interesting evidence of a further kind for "developmental conservation" of paths of adaption across phyla (Vargha-Khadem, Watkins, Alcock, Fletcher, & Passingham, 1995; Enard et al., 2002; Haesler et al., 2004). This kind of important, facilitatory change for rapid motor production and perhaps perception points out the necessary caution that an argument highlighting the social, motivational aspects of communication systems does not in any way preclude other alterations making vocal (or other) communication more feasible, including the speed-of-processing aspect hinted at in the FOXP2 expression work, or the requirement for adequate memory stores in both bird song and language.

What has changed in the human brain permitting language acquisition?

In this paper, we have attempted to draw together selected types of evidence about the organization and evolution of systems of social communication to seat human language learning in this context, ignoring the voluminous literature about the sensory, analytical, cognitive and linguistic nature of language. We have reviewed evidence to suggest the elaborate interconnection of social and motivational circuitry, in neuroanatomical terms. Current evidence further suggests an unusual combinatorial capacity in motivational systems over evolutionary time (Goodson, 2005). We suggest that the new linkage that has occurred in the evolution of development has been made between the neural representation of central caregivers (as to a lamb, certain mate or to a zebra finch "tutor"), the motivational systems, and cortical vocal and gestural learning systems. In species that have evolved vocal learning, influencing the desired individual's attention through behaviour, particularly vocalization, as indicated by gaze, contact or localization produces the most profound reward. A learning system, cortical in origin, of unusual power, has forged a unique link to the social circuit whereby social motivation "gates" vocal learning (Figure 3). Fig 3a diagrams such a proposed "gate" in the vole social behaviour network whereby dopaminergic reward systems, through the Nucleus Accumbens, become associated with specific individuals – as in the case of monogamous voles – in concert with VIaR receptor upregulation in the Ventral Pallidum), while Figure 3b hypothesizes a similar link or gate between recognition of a parent and a vocal learning system.

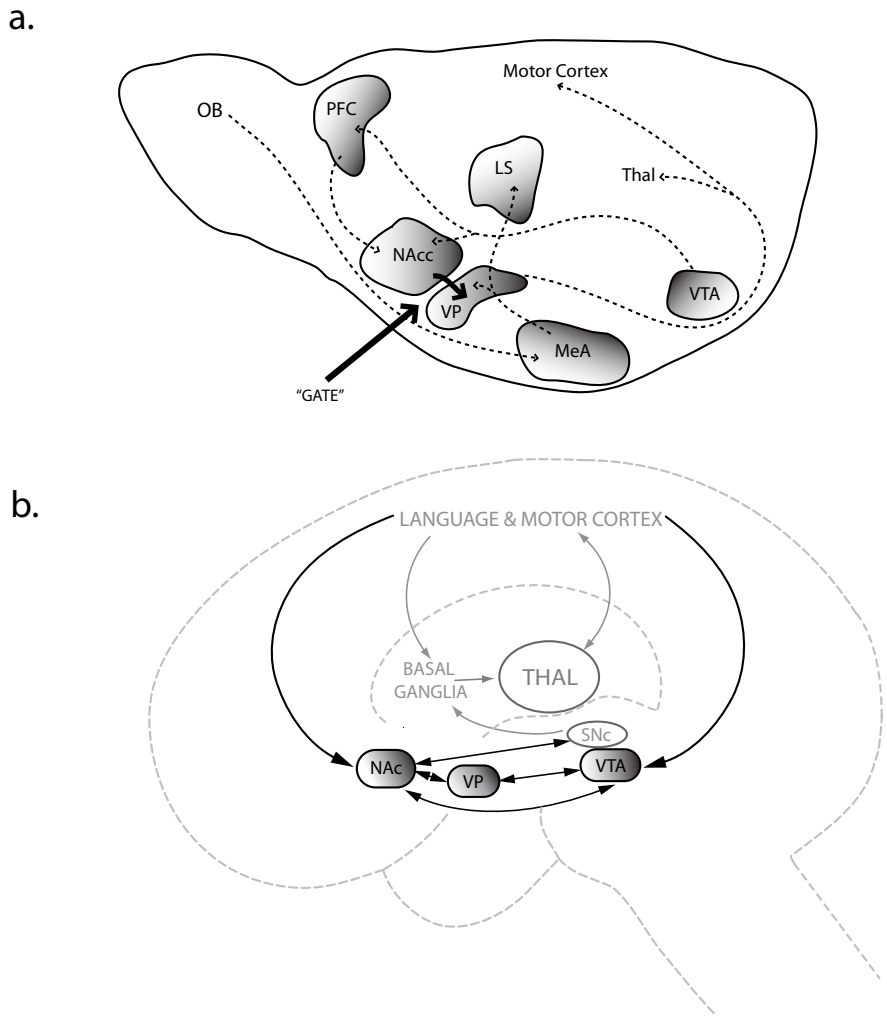

Figure 2.3 The social core of the language learning:

3a) Schematic of a rodent (vole) brain showing the social behavior network. The connection between the Nucleus Accumbens (reward center) and the Ventral Pallidum is highlighted as a potential "gate" for social motivation whereby the recognition of an individual conspecific becomes potentiated – for example, via the up-regulation of V1aR receptors in monogamous voles - modified from Young and Wang, 2004.

3b) Schematic of a human brain showing the basic motor circuitry of language in grey and its interconnectedness with reward centers that might serve similar social gating functions in humans. Abbreviations: LS – Lateral Septum, MeA – Medial Amygdala, NAc – Nucleus Accumbens, OB - Olfactory Bulb, PFC – Prefrontal Cortex, VP – Ventral Pallidum, VTA – Ventral Tegmental Area, SNc – Substantia Nigra pars caudalis, Thal – Thalamus.

A whole constellation of well-studied recognition systems stand in place to organize the recognition of those significant individuals who engage this newly connected, socially-gated nervous system. Faces are special; a number of avenues suggest special evolution of both the morphology and the recognition of the significance of eye gaze and conformation (Senju & Johnson, 2009); recognition of the mother's voice occurs through learning *in utero* (DeCasper & Fifer, 1980); the preferences of infants for familiar tastes, smell and contact are well known.

Implications

The essential feature of the argument is that language emerges from a newly forged linkage of motivational systems and a desire to influence an object of much desire, in the presence of a powerful learning device. We have some evidence from early types of brain disorders that while near-catastrophic damage to sensory and cortical systems allows preservation of language, alterations to disposition or motivation are disproportionately damaging. For principally technical reasons at this point, direct evidence for disproportional activation of motivational systems during early language learning in infancy, such as that described in the activation of immediate early gene systems for the preferred partner in monogamous voles, cannot presently be gathered in human infants. Given the rate of technological advances in this area, we expect that this situation will not remain so for long, as this recent study using near infrared spectroscopy to examine the activation of the cortex in young infants in social situations attests (Lloyd-Fox et al., 2009). In the near future we suggest that researchers in early language development turn their attention from the storage device, the cortex, to the basal forebrain and striatum, which provide the motivational structure for behavior.

A final cautionary note, which may in fact be a point of leverage for future research, is

that the powerful social circuitry for particular behaviours we have described is rarely stable for the lifetime of the individual. Both parenthood and childhood pass, and the nature of those attachments mutate and special learning epochs disappear. In animal models, the neural signatures of those recognition, attachment and learning systems pass too. In fact, motivational coupling may entirely reverse, for example, when a formerly loved home and family become noxious to newly mature animal seeking a new home turf. Learning a new vocabulary word in adolescence is rarely an orgiastic experience. Perhaps, however, we feel an echo of that earlier state when as adults we make ourselves understood in a new country, in a new language for the first time.

REFERENCES

Adar, E., Lotem, A. and Barnea, A. (2008). The effect of social environment on singing behavior in the zebra finch (Taeniopygia guttata) and its implication for neuronal recruitment. *Behav. Brain. Res*, 187(1), 178-184

Adret, P. (1993). Operant conditioning, song learning and imprinting to taped song in the zebra finch. *Animal Behavior*. 46, 149–59.

Alcock, J. (2009). *Animal Behavior: An Evolutionary Approach* (9th ed.). New York: Sinauer Associates.

Baptista, L.F., & Petrinovich, L. (1984). Social interaction, sensitive phases and the song template hypothesis in the white-crowned sparrow. Animal Behaviour, 32, 172–181.

Bates, E., Reilly, J., Wulfeck, B., Dronkers, N., Opie, M., Fenson, J., et al. (2001). Differential effects of unilateral lesions on language production in children and adults. *Brain and Language, 79*(2), 223-265.

Bates, E. A. (2004). Explaining and interpreting deficits in language development across clinical groups: Where do we go from here? *Brain and Language, 88*(2), 248-253.

Bellugi, U., Lichtenberger, L., Jones, W., Lai, Z., St. George, M. I. (2000). The neurocognitive profile of Williams syndrome: a complex pattern of strengths and weaknesses. *Journal of Cognitive Neuroscience*, 12 [Suppl 1], 7–29.

Biben, M. and Symmes, D. (1991). Playback studies of affiliative vocalizing in captive squirrel monkeys: Familiarity as a cue to response. *Behaviour,* 117:1-19.

Bigelow, A. E. (1998) Infants' sensitivity to familiar contingencies in social interaction. Infant Behavior and Development, 21, 149-162.

Bigelow, A. E., & Rochat, P. (2006) Two-month-old infants' sensitivity to social contingency in mother-infant and stranger-infant interaction, Infancy, 9(3), 313-325.

Bird, G., Catmur, C., Silani, G., Frith, C., Frith, U. (2006). Attention does not modulate neural responses to social stimuli in autism spectrum disorders. *NeuroImage* 31(4), 1614-1624.

Bowles, S. (2006). Group competition, reproductive leveling, and the evolution of human altruism. *Science, 314*(5805), 1569-1572.

Bolhuis, J. J., and Gahr, M. (2006). Neural mechanisms of birdsong memory. *Nature Reviews Neuroscience, 7*(5), 347-357.

Bolhuis, J. J., and Wynne, C. D. L. (2009) Can evolution explain how minds work? *Nature*, 458, 832-833.

Blair, R. J., Frith, U., Smith, N., Abell, F., and Cipolotti, L. (2002) Fractionation of visual memory: agency detection and its impairment in autism, *Neuropsychologia* 40 (2002), pp. 108–118.

Bloom, K., Russell, A., & Wessenberg, K. (1987).Turn taking affects the quality of infant vocalizations. *Journal of Child Language*, 14, 211-227.

Broad, K. D., Curley, J. P., and Keverne, E. B. (2006). Mother-infant bonding and the evolution of mammalian social relationships. *Philosophical Transactions of the Royal Society of London B, 361*, 2199-2214.

Cardinal, R. N., Parkinson, J. A., Hall, J., and Everitt, B. J. (2002). Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. *Neurosci Biobehav Rev, 26*(3), 321-352

Cheng, M. F. (2003). Vocal self-stimulation: From the ring dove story to emotion-based vocal communication. *Adv. Study Behav.*, 33:309-353.

Cheng, M. F. and Durand, S.E. (2004). Song and the Limbic Brain: A New Function for the Bird's Own Song. *Ann. NY Acad. Sci.,* 1016:611-627.

Chomsky, N. (1981). *Rules and Representations*. Columbia Univ. Press, New York.

Cohen, J., and Cheng, M. F. (1981) The role of the midbrain in courtship behavior of the female ring dove (*Streptopelia risoria*): evidence from radiofrequency lesion and hormone implant studies. *Brain Res.*, 207, 279–301.

Cohen, M. X., Schoene-Bake, J. C., Elger, C. E., and Weber, B. (2009). Connectivity-based

segregation of the human striatum predicts personality characteristics. *Nat Neurosci, 12*(1), 32-34.

Conboy, B. T., Brooks, R., Meltzoff, A. N., Kuhl, P. K. (2008). Joint engagement with language tutors predicts brain and behavioral responses to second-language phonetic stimuli. Pr*esented at ICIS, 1th,* Vancouver.

Cosmides, L & Tooby J (1992) Cognitive adaptations for social exchange. *In The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, eds Barkow J.H., Cosmides, L, Tooby, J. (Oxford Univ Press, New York), pp 163-228.

Damasio, H., Grabowski, T., Frank, R., Galaburda, A. M., and Damasio, A. R. (1994). The return of Phineas Gage: clues about the brain from the skull of a famous patient. *Science, 264*(5162), 1102-1105.

DeCasper, A. J., and Fifer, W. P. (1980). Of human bonding: newborns prefer their mothers' voices. *Science, 208*(4448), 1174-1176.

Depue, R. A., Morrone-Strupinsky, J. V. (2005) A neurobehavioral model of affiliative bonding: Implications for conceptualizing a human trait of affiliation. *Behavioral and Brain Sciences*. 28: 313–395.

Doupe, A. J. and Kuhl, P. K. (1999). Birdsong and human speech: common themes and mechanisms. *Ann. Rev. Neurosci*., 22, 567-631.

Dunbar, R.I.M and Shultz, S. (2007) Understanding primate brain evolution. *Phil. Trans. R. Soc. B* 362, 649-658.

Eales, L. A. (1989). The influences of visual and vocal interaction on song learning in zebra finches. *Anim. Behav*., 37,507–8.

Edinger, L. (Translation from German) *Investigations on the Comparative Anatomy of the Brain* Volumes **1–5** (Moritz Diesterweg, Frankfurt/Main, 1888–1903).

Enard, W., Przeworski, M., Fisher, S. E., Lai, C. S. L., Wiebe, V., Kitano, T., Monaco, A. P. & Paabo, S. 2002 Molecular evolution of FOXP2, a gene involved in speech and language. *Nature* **418**, 869-872.

Fernald, A. and Simon, T. (1984). Expanded intonation contours in mother's speech to newborns. *Dev. Psychol.,* 20,106-112.

Fernald, A., Kermanschachi, N. and Lees, D. (1984). The rhythms and sounds of soothing: Maternal vestibular, tactile, and auditory stimulation and infant state. *Biennial International Conference on Infant Studies*.

Fernald, A., Taeschner, T., Dun, J., Papousek, M., de Boysson-Bardies, B., and Fukui, I.(1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *J. Child Lang.,* 16,477–501.

Fernald, A. (1992a). Maternal vocalisations to infants as biologically relevant signals: An evolutionary perspective. In J.H. Barkow, L. Cosmides, and J. Tooby (Eds.). *The adapted mind: Evolutionary psychology and the generation of culture*. Oxford: Oxford University Press.

Fernald, A. (1992b). Meaningful melodies in mothers' speech to infants. In Papousek, H., Jurgens, U., and Papousek, M. (Eds.), *Nonverbal vocal communication: Comparative and developmental approaches*. Cambridge: Cambridge University Press, 262-282.

Fromkin, V., Krashen, S., Curtis, S., Rigler, D., Rigler, M. (1974). The development of language in Genie: a case of language acquisition beyond the "critical period". *Brain Lang.,* 1, 81-107

Fodor, J. A. (1983). *The Modularity of Mind: An Essay on Faculty Psychology*. MIT Press, Cambridge, MA.

Goldman-Rakic, P. S. (1999). The physiological approach: functional architecture of working memory and disordered cognition in schizophrenia. *Biol Psychiatry, 46*(5), 650-661.

Goldstein, M. H., King, A. P., and West, M. J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proc. Nat. Acad. Sci.,* 100(13), 8030-8035.

Goldstein, M. H., and Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychol. Sci.,* 19(5), 515-523.

Goodson J. L. and Evans A. K. (2004). Neural responses to territorial challenge and nonsocial

stress in male song sparrows: segregation, integration, and modulation by a vasopressin V1 antagonist. *Horm. Behav.,* 46, 371–381.

Goodson, J. L. (2005). The vertebrate social behavior network: evolutionary themes and variations. *Hormones and Behavior, 48*, 11-22.

Goodson, J. L., Evans, A. K., Lindberg, L. and Allen, C. D. (2005). Neuro-evolutionary patterning of sociality. *Proc. R. Soc. B*, 272:227–235

Goodson, J. L. (2008). Nonapeptides and the evolutionary patterning of sociality. *Progress in Brain Research, 170*, 3-13.

Hara, E., Kubikova, L., Hessler, N. A., and Jarvis., E. D. (2007) Role of midbrain dopaminergic system in the modulation of vocal brain activation by social context. *European Journal of Neuroscience* 25, 3406-3416.

Hammock, E. A. D. and Young, L. J. (2004). Functional Microsatellite Polymorphism Associated with Divergent Social Structure in Vole Species. *Mol. Biol. Evol.,* 21(6):1057-1063.

Hammock, E. A. D. and Young, L. J. (2005). Microsatellite Instability Generates Diversity in Brain and Sociobehavioral traits. *Science,* 308(5728):1630-1634.

Happe, F. G. (1995) "The role of age and verbal ability in the theory of mind task performance of subjects with autism." Child Development 66(3): 843-55, 1995.

Haesler, S., Wada, K., Nshdejan, A., Morrisey, E. E., Lints, T., Jarvis, E. D. & Scharff, C. 2004 FoxP2 Expression in Avian Vocal Learners and Non-Learners. *Journal of Neuroscience* **24**, 3164-3175.

Hausberger, M., Richard-Yris, M., Henry, L., Lepage, L. and Schmidt, I. (1995). Song sharing reflects the social organization in a captive group of European starlings (*Sturnus vulgaris*). *J. Comp. Psychol.,* 109:222-241.

Hessler, N.A., and Doupe, A.J. (1999). Social context modulates singing-related neural activity in the songbird forebrain. *Nat. Neurosci.*, 2, 209-211.

Houx, B.B., & ten Cate, C. (1999) Learning from playback tutor song in zebra finches: is there an effect of operant contingency? *Animal Behaviour* **57**: 837–845

Huang, Y. C., and Hessler, N. A. (2008). Social modulation during songbird courtship potentiates midbrain dopaminergic neurons. *PLoS One, 3*(10), e3281.

Immelmann, K. (1969). Song development in the zebra finch and other estrildid finches. In Hinde, R. A. ed. *Bird Vocalizations*, Cambridge Univ. Press, London. pp. 61–74.

Insel, T. R., Wang, Z. and Ferris, C. F. (1994). Patterns of brain vasopressin receptor distribution associated with social organization in microtine rodents. *J. Neurosci.,* 14, 5381-5392.

Izuma, K., Saito, D. N., and Sadato, N (2008) Processing of Social and Monetary Rewards in the Human Striatum, *Neuron*, 58(2) 284-294.

Jarvis, E. D., Gunturkun, O., Bruce, L., Csillag, A., Karten, H., Kuenzel, W., et al. (2005). Avian brains and a new understanding of vertebrate brain evolution. *Nature Reviews Neuroscience, 6*(2), 151-159.

Jarvis, E. D. (2007). The Evolution of Vocal Learning Systems in Birds and Humans. In: In Jon H. Kaas ed. *Evolution of Nervous Systems* Academic Press, New York, Vol. 2, pp 213-228.

Johnston, R. E. (2005). Communication by mosaic signals: Individual recognition and underlying neural mechanisms. In *Chemical Signals in Vertebrates* (Vol. 10, pp. 269-282).

Jones, W., Bellugi, U., Lai, Z., et al. (2000). II. Hypersociability in Williams syndrome. Journal of Cognitive Neuroscience, 12(Suppl 1), 30–46.

Kirschner, M. W. & Gerhart, J. C. (2005) *The Plausibility of Life: Resolving Darwin's Dilema*. New HAven, CT: Yale University Press.

Klinger, L., Dawson, G., and Renner, P. (2002). Autistic disorder. In E. Mash and R. Barkley (Eds.), *Child* psychopathology (2nd edn.), New York: Guilford Press. pp. 409–454.

Kuhl P. K., Tsao F. M., Liu H. M., 2003. Foreign-language experience in infancy: effects of short-term exposure and social interaction on phonetic learning. *Proc. Natl. Acad. Sci.* USA 100: 9096–101.

Kuhl, P. K., Coffey-Corina, S., Padden, D. and Dawson, G. (2005). Links between social and linguistic processing of speech in preschool children with autism: Behavioral and electrophysiological evidence. *Dev. Sci.*, 8,1-12.

Kuhl PK. 2007. Is speech learning 'gated' by the social brain? *Dev. Sci.* 10: 110–20.

Laing, E., Butterworth, G., Ansari, D., et al. (2002). Atypical development of language and social communication in toddlers with Williams syndrome. Developmental Science, 5(2), 233–46.

Lane, H. L. (1976). *The Wild Boy of Aveyron*. Harvard Univ. Press, Cambridge, MA.

Lebreton, M., Barnes, A., Miettunen, J., Peltonen, L., Ridler, K., Veijola, J., et al. (2009). The brain structural disposition to social interaction. *Eur J Neurosci*.

Lim, M. M. and Young, L. J. (2002). Blockade of vasopressin V1a receptors in the ventral pallidum prevents partner preference formation in monogamous prairie voles. *Society for Neuroscience Abstract; 2002*. Orlando, FL: 32nd Annual Meeting: 89.2.

Lim, M. M., Murphy, A. Z. and Young, L. J. (2004). Ventral striato-pallidal oxytocin and vasopressin V1a receptors in the monogamous prairie vole (Microtus ochrogaster). *J. Comp. Neurol.*, 468,555–570.

Lim, M. M. and Young, L. J. (2004) Vasopressin-dependent neural circuits underlying pair bond formation in the monogamous prairie vole. *Neuroscience*, 125,35-45.

Lim, M. M., Hammock, E. A. D., Young, L. J. (2004). The role of vasopressin in the genetic and neural regulation of monogamy. *J. Neuroendocrinol.*, 16(4),325-332.

Lipkind, D., Nottebohm, F., Rado, R., Barnea, A. (2002) Social change affects the survival of new neurons in the forebrain of adult songbirds. *Behav. Brain Res.*, 133,31-43.

Lloyd-Fox, S., Blasi, A., Volein, A., Everdell, N., Elwell, C. E., Johnson, M. H., (2009). "Social

perception in infancy: A near infrared spectroscopy study." Child Development 80(4), 986-999.

Locke J. (1993). *The Path to Spoken Language*. Harvard Univ. Press, Cambridge, MA.

Mann, N. I. and Slater, P. J. B. (1995). Song tutor choice by zebra finches in aviaries. *Animal Behav*., 49(3):811-820.

Martinez-Garcia, F., Martinez-Marcos, A., and Lanuza, E. (2002). "The pallial amygdala of amniote vertebrates: evolution of the concept, evolution of the structure." *Brain Res Bull* 57(3-4): 463-9.

Martinez-Garcia, F., Novejarque, A., and Lanuza, E. (2007). Evolution of the Amygdala in Vertebrates. In Jon H. Kaas ed. *Evolution of Nervous Systems*. Academic Press, New York. pp: 255-334.

Medina, L (2007). Do Birds and Reptiles Posess Homologues of Mammalian Visual, Somatosensory and Motor Cortices? In Jon H. Kaas ed. *Evolution of Nervous Systems*. Academic Press, New York. Vol 2 pp: 163-194.

Mogenson, G. J., Jones, D. L., and Yim, C. Y. (1980). From motivation to action: functional interface between the limbic system and the motor system. *Prog Neurobiol, 14*(2-3), 69-97.

Moll, H., and Tomasello, M. (2007). Cooperation and human cognition: the Vygotskian intelligence hypothesis. *Philosophical Transactions of the Royal Society of London B, 10*.

Morgane, P.J., Galler, J.R., and Mokler, D.J., (2005). A review of systems and networks of the limbic forebrain/limbic midbrain. Progress in Neurobiology, 75: 143-160

Marler P. (1970). A comparative approach to vocal learning: song development in white-crowned sparrows. *J. Comp. Physiol. Psychol*., 71:1–25

Marler, P. and Sherman, V. (1985). Innate differences in singing behavior of sparrows reared in isolation from adult conspecific song. *Animal Behav*., 33:57–71

Nacewicz, B. M., Dalton, K. M., Johnstone, T., Long, M. T., McAuliff, E. M.,Oakes, T. R.,

Alexander, A. L. and Davidson, R. J. (2006). Amygdala volume and nonverbal social impairment in adolescent and adult males with autism. *Arch. Gen. Psychiatry*, 63:1417-1428.

Neumann D, Spezio ML, Piven J, Adolphs R (2006) Looking you in the mouth: abnormal gaze in autism resulting from impaired top-down modulation of visual attention. Soc Cogn Affect Neurosci, 1(3), 194-202

Newman, S. W. (1999). The medial extended amygdala in male reproductive behavior. A node in the mammalian social behavior network. *Annals of the New York Academy of Sciences, 877*, 242-257.

Nottebohm, F. 1976. Neural lateralization of vocal control in a passerine bird. *J. Exp. Zool*. 177, 229–262.

Nowicki, S. (1989). Vocal plasticity in captive black-capped chickadees: The acoustic basis of call convergence. *Animal Behav.,* 37: 64-73.

Osterling, J., and Dawson, G. (1994). Early recognition of children with autism: a study of first birthday home videotapes. *J. Autism and Dev. Disorders*, 24:247-257.

Osterling, J., Dawson, G., and Munson, J. (2002). Early recognition of 1-year-old infants with autism spectrum disorder versus mental retardation. *Dev. Psychopathol.,* 14:239–251.

Owren M.J., Dieter J.A., Seyfarth R.M. and Cheney D.L. (1993). Vocalizations of rhesus (*Macaca mulatta*) and Japanese (*M. fuscata*) macaques cross-fostered between species show evidence of only limited modification. *Dev. Psychobiol*., 26:389-406.

Pagel, M. (2009). Human language as a culturally transmitted replicator. *Nat Rev Genet, 10*(6), 405-415.

Papousek, M., Papousek, H., and Bornstein, M. H. (1985). The naturalistic vocal environment of young infants: On the significance of homogeniety and variability in parental speech. In Field, T. and Fox, N. Eds. *Social perception in infants*. [pp 269 -297] Ablex, Norwood, NJ

Paul, B. M., Snyder, A. Z., Haist, F., Raichle, M. E., Bellugi, U., and Stiles, J. (2009) Amygdala response to faces parallels social behavior in Williams syndrome *Soc Cogn Affect*

Neurosci. 4(3): 278-285

Pepperberg, I. M. (1999). *The Alex Studies: Cognitive and Communicative Abilities of Grey Parrots*. Harvard University Press, Cambridge, MA.

Preston, S. D., and de Waal, F. B. M. (2001). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences, 25*(01), 1-20.

Puelles, L. (2001). Thoughts on the development, structure and evolution of the mammalian and avian telencephalic pallium. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 356*(1414), 1583-1598.

Puelles, E. Kuwana, E. Puelles, A. Bulfone, K. Shimamura, J. Keleher, Smiga, S., and Rubenstein, J.L.R. (2000) Pallial and subpallial derivatives in the embryonic chick and mouse telencephalon, traced by the expression of the genes Dlx-2, Emx-1, Nkx-2.1, Pax-6, and Tbr-1. *J. Comp. Neurol*. 424, pp. 409–438

Reilly J, Losh M, Bellugi U, and Wulfeck, B. (2004) "Frog where are you?" Narratives in children with specific language impairment, early focal brain injury and Williams syndrome. *Brain Lang* 88:229 –247.

Richerson, P. J., and Boyd, R. (2005). *Not by Genes Alone. How Culture Transformed Human Evolution*. Chicago, IL: Univ. of Chicago Press.

Roush, R. S. and Snowdon, C. T. (1994). Ontogeny of food-associated calls in cotton-top tamarins. *Animal Behav.,* 47, 263-273.

Roush, R. S. and Snowdon, C. T. (1999). The effects of social status on food associated calling behavior in captive cotton-top tamarins. *Animal Behav*, 58(6),1299-1305.

Sasaki, A., Sotnikova, T.D., Gainetdinov, R.R. and Jarvis, E.D. (2006). Social context-dependent singing-regulated dopamine. *J. Neurosci.,* 26, 9010-9014.

Schumann, C. M., Hamstra, J., Goodlin-Jones, B. L., Lotspeich, L. J., Kwon, H., Buono- core, M. H., Lammers, C. R., Reiss, A. L. and Amaral, D. G. (2004). The amygdala is enlarged in children but not adolescents with autism; the hippocampus is enlarged at all ages. *J. Neurosci.,* 24, 6392– 6401.

Senju, A., and Johnson, M. H. (2009). The eye contact effect: mechanisms and development. *Trends in Cognitive Sciences, 13*(3), 127-134.

Senju, A., Southgate, V., White, S., & Frith, U. (2009). Mindblind eyes: An absence of spontaneous theory of mind in Asperger syndrome. *Science*, 325, 883-885.

Seyfarth, R.M. and Cheney, D. L. (1986). Vocal development in vervet monkeys. *Animal Behav*., 34, 1640–1658.

Smith, H. J., Newman, J. D., Hoffman, H. J. and Fetterley, K. (1982a). Statistical discrimination among vocalizations of individual squirrel monkeys (*Sairniri sciureus*). *Folia Primatol*., 37, 267-279.

Smolker, R.A. and Pepper, J.W. (1999). Whistle convergence among allied male bottlenose dolphins (Delphinidae, Tursiops sp.). *Ethology*, 105, 595-617.

Snowdon, C. T. and Elowson, A.M. (1994) Pygmy Marmosets, *Cebuella pygmaea*, modify vocal structure in response to changed social environment. *Animal Behav*., 47, 1267-1277

Snowdon, C. T. and Elowson, A.M., and Roush, R. S. (1997) Social influences on vocal development in New World Primates. In C.T. Snowdon and M. Hausberger, eds. *Social influences on vocal development*. Cambridge University Press, Cambridge, pp 234-248.

Stern, D. N., Spieker, S. and MacKain, K. (1982). Intonation contours as signals in maternal speech to prelinguistic infants. *Dev. psychol*., 18(5), 727-749.

Striedter, G. F. (1997). "The telencephalon of tetrapods in evolution." *s* Evol 49(4), 179-213.

Striedter, G. (2005). *Principles of Brain Evolution*. New York: Sinauer.

Thompson, P. M., Lee, A. D., Dutton, R. A., Geaga, J. A., Hayashi, K. M., Eckert, M. A., Bellugi, U., Galaburda, A. M., Korenberg, J. R., Mills, D. L., Toga, A. W., and Reiss, A. L. (2005). Abnormal cortical complexity and thickness profiles mapped in williams syndrome. *J. Neurosci.,* 25(16):4146-4158.

Thorpe, W. H. (1958). The learning of song patterns by birds with especial reference to the song of the chaffinch, *Fringilla coelebs*. *Ibis*, 100:535–70.

Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. harvard: Harvard University Press.

Trezza, V., Damsteegt, R., Marijke Achterberg, E. J., and Vanderschuren, J. M. J., Louk (2011) Nucleus Accumbens μ-Opioid Receptors Mediate Social Reward *The Journal of Neuroscience*, 31(17): 6362-6370.

Vargha-Khadem, F., Watkins, K., Alcock, K., Fletcher, P. & Passingham, R. 1995 Praxic and nonverbal cognitive deficits in a large family with a genetically transmitted speech and language disorder. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 930-933.

Warneken, F., and Tomasello, M. (2008). The roots of human altruism. *British Journal of Psychology, 00*(1-18).

Wang, Z., Zhou, L., Hulihan, T. J. and Insel, T. R. (1996). Immunoreactivity of central vasopressin and oxytocin pathways in microtine rodents: a quantitative comparative study. *J. Comp. Neurol.*, 366:726-737.

West, M. J., and King, A. P. (1988). Female visual displays affect the development of male song in the cowbird. *Nature*, 334(6179):244-246.

Wild, J. M. (1993). Descending projections of the songbird nucleus robustus archistriatalis. *J. Comp. Neurol.,* 338, 225-241.

Wild, J. M., Karten, H. J. and Frost, B. J. (1993) Connections of the auditory forebrain in the pigeon (*Columba livia*). *J. Comp. Neurol.,* 337, 32–62.

Wilson, D. S., Van Vugt, M., and O'Gorman, R. (2008). Multilevel selection theory and major evolutionary transitions: implications for psychological science. *Current Directions in Psychological Science, 17*(1), 6-9.

Winter, P., Handley, P., Ploog, D., and Schott, D. (1973). Ontogeny of squirrel monkey calls under normal conditions and under acoustic isolation. *Behaviour,* 47, 230-239.

Wong Fillmore, L. (1991). When learning a second language means losing the first. *Early Childhood Res. Quar.*, 6:323-346.

Young, L. J., Nilsen, R., Waymire, K. G., MacGregor, G. R. and Insel, T. R. (1999). Increased affiliative response to vasopressin in mice expressing the V1a receptor from a monogamous vole. *Nature,* 400,766-768.

Young, L. J., and Wang, Z. (2004). The neurobiology of pair bonding. *Nature neuroscience, 7*(10), 1048-1054.

CHAPTER 3

SOCIALLY GUIDED STATISTICAL LEARNING: HOW INFANTS LEARN TO TALK[3]

Abstract

Babbling is a ubiquitous vocal behavior of infants during the first year, but its function is poorly understood. Early non-cry vocalizations are usually considered an epiphenomenon of motor practice. We present the first evidence that babbling has an important function in learning new phonological patterns. Manipulations of caregivers' responses to babbling reveal the ability of nine-month-old infants to learn new patterns of vocalizing from social and statistical regularities. Infants learn new vocal forms only when caregivers' reactions are contingent on babbling and contain phonologically-varied exemplars. Infants do not learn from noncontingent responses nor from repetitive speech. Our findings support a new mechanism of phonological development, socially guided statistical learning, that is responsible for the emergence of language-specific phonological patterns.

Introduction

A prerequisite for fluent speech is the ability to perceive and produce the phonological patterns of the ambient language. Most research on early language development focuses on speech perception, as prelinguistic vocal productions (babbling) are typically considered an product of immature articulatory control *(*Kent, 1981). However, in songbirds, long considered a

[3] (Goldstein, M. H., Syal, S., and Schwade, J. A. manuscript under review)

model for human vocal learning and communication (Doupe & Kuhl, 1999), immature vocal productions have important functional significance. Young songbirds learn from the reactions of older conspecifics to their early vocalizations (West & King, 1988).

Like songbirds, human infants exhibit substantial developmental changes in vocal production. Infants produce non-cry vocalizations in the first days of life but do not attain their first words until approximately 12 months of age. Experience clearly plays a role in shaping prelinguistic vocalizations towards the sound patterns of the ambient language (Vihman, 1993). By approximately 8 months of age, several suprasegmental features of vocalizing have come to resemble those of the ambient language (Levitt & Wang, 1991; Whalen, Levitt, & Goldstein, 2007). By 9-10 months infants begin to produce mature syllables (Oller, 2000) and structure sequences of syllables in accordance with the phonological rules of their language environment (Boysson-Bardies, 1993; Boysson-Bardies & Vihman, 1991). For example, Nigerian infants typically learn Yoruba, a language with many vowel-initial words. The babbling of these 10 – 12-month-old infants has a typical disyllable structure of vowel-consonant-vowel (VCV), in contrast to the typical CVCV disyllable pattern found in the babbling of French and English infants (Boysson-Bardies, 1993).

What mechanisms drive language-specific changes in babbling? The processes of change that drive vocal development are not well understood. Most work focuses on the maturation of the vocal tract, but comparative studies of vocal development in songbirds found that social learning plays an important role in developmental change. For example, young male cowbirds (*Molothrus ater*) rely on the reactions of females to shape their immature sounds into functional song (West & King, 1988; King, West, & Goldstein, 2005). Based on the avian work, our previous research has shown that prelinguistic infants use social feedback that is contingent on

their babbling to guide their productions towards well-formed speech (e.g., fully-resonant vowels and canonical syllables, Goldstein, King, & West, 2003). In addition, contingent social feedback can facilitate the production of specific phonological patterns such as V or CV (Goldstein & Schwade, 2008).

In the present study, we hypothesized that contingent feedback, in conjunction with the statistical variability of exemplars in caregivers' speech, would promote learning of a rare or novel phonological form. In multiple domains, increased variability in exemplars facilitates generalization of learning in infants and adults (Gómez, 2002; Tenenbaum & Griffiths, 2001). Thus we manipulated both the timing and the variability of caregivers' verbal responses to babbling. Caregiver-infant dyads participated in two 30-minute play sessions, scheduled 24 hours apart. The first session familiarized infants and their caregivers with the playroom. The second session had an ABA design and was divided into three 10-minute periods: Baseline-1, Social Response, and Baseline-2. During each Baseline period, caregivers were asked to play as they would at home. During the Social Response period, the experimenter gave instructions to caregivers via wireless headphones to control the form and timing of their responses to infant vocalizations.

Dyads (n = 60) participated in one of two contingent social response conditions, in which the experimenter cued caregivers to respond to each infant vocalization (except vegetative and fussing sounds). Caregivers responded by speaking a Yoruban (VCV-structured) word while moving closer to, smiling at, and touching their infants. The specific Yoruban word given to the caregivers over the wireless link served as the cue to respond. In the Contingent-Variable (C-Var) group, caregivers said a different Yoruban word during each response, e.g. [ɛnu], [idi], [abo], [olu], [ika]. In the Contingent-Repetition (C-Rep) group, caregivers repeated the same

Yoruban word during each response.

Materials and Methods

Participants. Sixty caregiver-infant dyads were tested (mean infant age 9 months, 19 days; range 8 months 4 days to 10 months 26 days). Families were recruited from birth announcements printed in local newspapers and from advertisements. All infants included in the final sample were exposed to English as their primary language. An additional 8 infants were tested but excluded for the following reasons: infants cried or were excessively fussy (n = 5), infants produced fewer than 5 vocalizations in any experimental period (n = 1), equipment failure (n = 1), and failure to return for a second session (n = 1). Participants were given an infant t-shirt or bib for their participation. Twenty infants were tested in each of the three (C-Var, C-Rep, YC) conditions. Infant gender was counterbalanced across conditions.

Apparatus. Infants were tested in large laboratory playroom (3.7 m x 5.5 m) containing a toybox, play mat, and infant toys. Thus infants were not forced to interact with their caregivers; they could become distal and explore the toys in the room. Sessions were videorecorded via three remotely-controlled wall-mounted digital cameras routed to a video mixer. To obtain high-quality recordings of infant vocalizations, each infant wore a pair of denim overalls that concealed a microphone and small wireless transmitter. Each caregiver wore a wireless microphone on her collar with the transmitter concealed in a pouch worn at his/her waist. During the second session, the experimenter instructed caregivers via wireless headphones.

Stimuli. To obtain accurate exemplars of the speech stimuli, we recorded a native speaker of Yoruban speaking a list of VCV-structured Yoruban words. At the beginning of the second session, a CD player was used to play this list of Yoruban words for all caregivers. In the C-Rep condition, we used one of the three Yoruban words most representative of the full list in terms of

the vowels and consonants that they contained (i.e. [agə], [etɪ], [ɛnu]). The three words were then counterbalanced across infants. A Macintosh Powerbook was used to play Quicktime sound files of the Yoruban words to cue the caregivers.

Yoked control procedure. In the YC condition, the form and timing of caregivers' responses were yoked to those of a randomly-selected C-Var caregiver[12]. To control the timing of YC caregivers' responses, we created a CD track of each C-Var caregivers' contingent utterances. Extraneous sounds such as infant vocalizations and toy noises were removed while preserving the timing of caregivers' speech. During the social response period, these CD tracks were played back to the yoked control caregivers over the wireless headphones. Yoked control caregivers were asked to repeat the vocalizations that they heard over their headphones while getting closer to, touching, and smiling at their infants. Thus the YC caregivers responded using the same Yoruban words and timing as the C-Var caregivers with whom they were paired, and their infants received the same amount and type of social feedback as the C-Var infants, but the feedback was not synchronized with their vocalizations.

Vocalization coding. Infant vocalizations were categorized by one of 7 coders, who were first trained until they exceeded 90% reliability with the first author. To assess the phonological patterns in infants' babbling, we tallied each infant's vocalizations in each experimental period. Each vocalization that contained consonants and vowels was identified as beginning with a consonant or with a vowel. Vocalizations that began with a consonant were coded as having English phonological structure (EPS); vocalizations that began with a vowel were coded as having Yoruban phonological structure (YPS). The number of CV-structured syllables of each type was calculated separately for each infant in each experimental period. Analyses were conducted on the proportion of syllables with YPS to total number of syllables with CV-

structure. To assess reliability, a second coder independently re-scored the phonological patterns of infants' vocalizations for 20% of sessions. Reliability in classification of infant vocalizations was .96 (range .87- 1.00).

Measuring vocal imitation. To calculate matching during the social response period, we phonetically transcribed infant vocalizations and caregivers' utterances using the International Phonetic Alphabet. Transcriptions were made by one of four coders who had received extensive training with the IPA. Reliability was calculated for 20% of sessions (mean reliability = .92, range .84 - .96). From the transcripts of the social response period, we counted the number of infant phones in each vocalization. We then compared each infant utterance to the caregiver's previous utterance. We counted the number of infant phones that matched any phoneme in the caregiver's previous utterance, divided by the number of phones in the infant utterance. When calculating matching scores, we gave infants credit for partial matches. Thus, if a caregiver said [aba], and the infant's next utterance was [bu], the infant received credit for producing a match in one of his two phones (match score = 50% for that utterance). Each infant received a mean matching score that averaged matching scores across all utterances in the social response period.

Results

To assess the effects of input variability on vocal learning, we conducted a 2 (Condition: C-Var, C-Rep) x 2 (Period: Baseline-1, Social Response) mixed ANOVA on the proportion of CV-structured syllables with Yoruban (VCV) phonological structure (YPS) to total number of CV-structured syllables. There was a significant interaction of Condition x Period ($F_{1,38} = 6.44$, $P = .02$, $\eta_p^2 = .15$). We followed up the interaction with tests of simple main effects on condition. For infants in the C-Var condition, there was a significant effect of period on

proportion of YPS vocalizations ($F_{1,19} = 16.70$, $P = .001$, $\eta_p^2 = .47$, Figure 1). C-Var infants increased their production of YPS vocalizations from Baseline-1 to Social Response. A significant number (14 out of 20) of C-Var infants followed this pattern, Wilcoxon signed-rank $z = -3.35$, $P = .001$. In contrast, infants in the C-Rep condition did not show significant changes in proportion of YPS vocalizations across periods ($F_{1,19} = .15$, $P = .71$). Thus only the C-Var infants showed learning of the new phonological pattern.

We then assessed the effects of contingency on learning by testing a group of infants in a yoked control (YC) condition. Input for YC infants was matched in form and timing to that of the C-Var group, but was not contingent on their vocalizations. To assess vocal learning, we conducted a 2 (Condition: C-Var, YC) x 2 (Period: Baseline-1, Social Response) mixed ANOVA on the proportion of CV-structured syllables with YPS to total number of CV-structured syllables. There was a significant interaction of Condition x Period ($F_{1,38} = 4.57$, $P = .04$, $\eta_p^2 = .11$, Figure 1). Tests of simple main effects on condition showed that infants in the C-Var condition significantly increased production of YPS vocalizations from Baseline-1 to Social Response. In contrast, YC infants did not show significant changes in proportion of YPS vocalizations across periods ($F_{1,19} = .01$, $P = .94$). Thus contingency is a necessary component of social feedback that facilitates learning. Additional analyses showed that the groups did not differ in initial level of production of YPS vocalizations, and the C-Var infants maintained their level of YPS production into the second Baseline period.

To assess whether differences in learning across conditions were due to pre-existing differences at baseline, we conducted a one-way ANOVA with condition as a factor on proportion of YPS vocalizations in Baseline-1. There was no significant main effect of condition, $F\ (2, 57) = .71$, $p = .50$, $\eta_p^2 = .02$. Thus the three groups had similar levels of YPS

vocalizations during Baseline-1.

To test the stability of YPS production over the three periods, we conducted one-way ANOVAs on infants' proportion of YPS vocalizations. For the C-Var infants, there was a significant main effect of period, $F\ (2, 38) = 5.67, p = .007, \eta_p^2 = .23$ (Figure 1). These infants increased their production of YPS from Baseline-1 to Social Response, Tukey's post-hoc test $p < .05$. There was no significant difference in production of YPS between Social Response and Baseline 2. For C-Rep infants, there was no significant effect of period, $F\ (2, 38) = 1.28, p = .29$. For YC infants, there was also no significant effect of period, $F\ (2, 38) = .004, p = .996$.

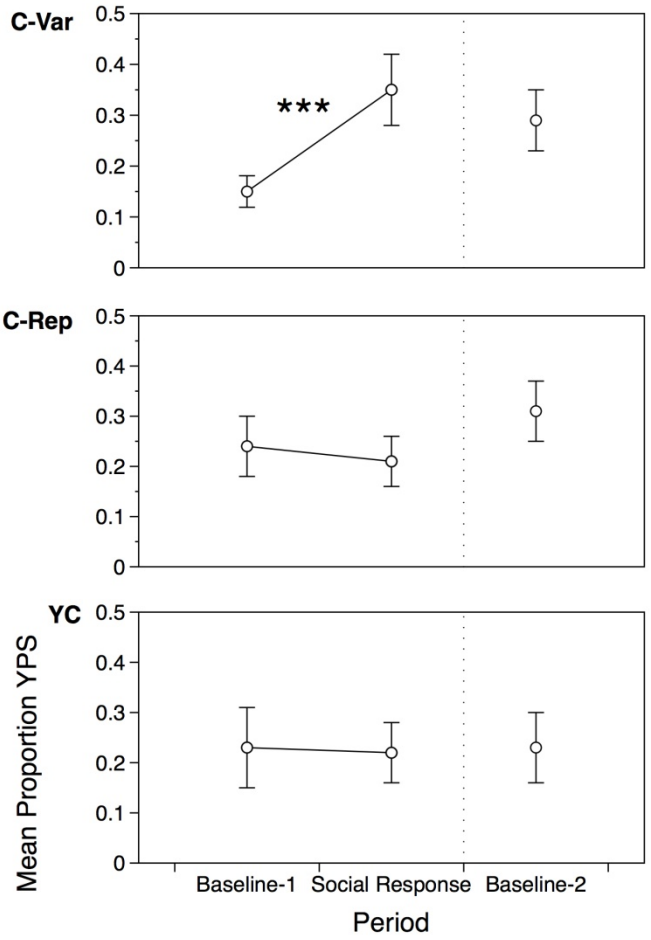

Figure 3.1. Mean proportion of syllables with Yoruban phonological structure (YPS) by condition. YPS syllables are vowel-consonant-vowel (VCV) in form. C-Var denotes infants who experienced contingent social feedback with variable speech. These infants received a different Yoruban word each time they vocalized. C-Rep denotes infants who received contingent feedback with repetitive speech.

They received the same Yoruban word each time they babbled. YC denotes yoked control infants. They received feedback that was identical to the C-Var group in amount and timing, but the feedback was contingent on their vocalizing. Error bars are s.e.m. *** $P = .001$.

Was the C-Var infants' increase in YPS vocalizations due to statistical learning and generalization of the Yoruban pattern, or to mimicry of their caregivers' speech? For each caregiver utterance during the Social Response period, we measured the amount of matching between the caregivers' phonemes and the phones of the infant vocalization that immediately followed. To compute a chance level of vocal matching, we shuffled the sequences of vocalizations to create a random distribution of pairs of caregiver-infant vocalizations. The observed amount of infant vocal mimicry (M = 15.10%, SD = 12.48%) was not significantly different from chance (M = 12.32%, SD = 12.22%, $t(18) = .70$, P = .49).

Discussion

Infants' production of a rare phonological structure was rapidly facilitated by social feedback that was contingent on babbling and contained phonemic variability. Only C-Var infants modified their vocal repertoires from the Baseline-1 to the Social Response periods, learning to reliably produce the Yoruban VCV form. These changes were maintained throughout the Baseline-2 period. Neither the C-Rep nor the YC infants changed their phonological repertoires, indicating that neither contingency nor exemplar variability alone was sufficient to promote learning. When C-Var infants produced Yoruban speech, they generalized the VCV pattern to new combinations of phonemes that were not present in the social feedback. For example, if the caregiver said [ada], the next infant vocalization might be [igi]. Infants thus picked up on phonological regularities in their caregivers' speech and used these patterns to guide their vocalizations. They did not mimic the same phonemes produced by their parents.

These data support socially guided statistical learning as a mechanism of prelinguistic vocal development and early language learning.

Our finding of non-imitative learning stands in contrast to previous reports of infants mimicking vowel sounds presented in isolation (Kuhl & Meltzoff, 1996; Legerstee, 1990). The difference between the present data and previous findings is not a question of the phonemic complexity of the exemplars, as previous research (Goldstein & Schwade, 2008) found that infants presented with a varying set of vowels also generalized to phonemes that did not match mothers' previous utterances. The key difference seems to be the amount of variability in the input. In general, across multiple domains, infants and adults show greater generalization of learning when given more variability in the training exemplars (Gómez, 2002; Tenenbaum & Griffiths, 2001; Rost & McMurray, 2009). Studies of artificial grammar learning show strong effects of variability on infants' recognition of patterns in the input. For example, increasing the variability of adjacent elements in an artificial grammar facilitates learning of nonadjacent dependencies (Gómez, 2002; Gómez & Maye, 2005).

The present study adds an important dimension to our understanding of language development beyond that provided by speech perception studies, which do not require contingent presentation for infants to successfully identify patterns of sounds (Gómez,2002; Marcus, Vijayan, Bandi Rao, & Vishton, 1999; Saffran & Aslin, 1996; Saffran & Thiessen, 2003). Why does contingency facilitate speech production but, seemingly, not perception? Social contingency is important for facilitating shared attention, which promotes learning (Kuhl, 2007; Meltzoff, Kuhl, Movellan, & Sejnowski, 2009; Movellan & Watson, 2002). In addition, contingency typically makes infants' and caregivers' utterances temporally proximate, which may facilitate infants' comparison of their own sounds with those of others (Goldstein & Schwade, 2008).

Because creating a new utterance that follows a phonological pattern is more difficult than recognizing an example of the pattern, contingency may be more useful for production than perception. However, the role of contingency in facilitating perception has received little attention, so there may be effects on perception as well.

Social contingency also enhances the reward value of early vocalizing (Goldstein, Schwade, & Bornstein, 2009) and points to the importance of subcortical brain regions such as the basal ganglia in learning language. For example, contingent feedback that activates reward-related striatal regions facilitates learning of non-native phonological contrasts in adults (Delgado, Nystrom, Fisell, Noll, & Fiez, 2000; Tricomi, Delgado, McCandliss, McClelland, & Fiez, 2006). An infant experiencing consistent feedback from caregivers for babbling will learn expectations for social events (e.g. adult speech) that follow vocalizing. These expectations may drive further learning about speech, as neural circuits underlying reward-based learning have been found to play a role in information-seeking (Bromberg-Martin & Hikosaka, 2009). The expectation of a reward activates the same neural regions – mid-brain dopamine neurons – as the reward itself (Bromberg-Martin & Hikosaka, 2009). The present findings implicate an important role of reward, prediction, and its associated subcortical circuitry in learning language.

These data are the first experimental evidence of prelinguistic vocal learning from distributional input. The findings support socially guided statistical learning as a crucial mechanism of phonological development, providing a pathway for social interaction to influence language acquisition. The consistent reactions of caregivers to babbling (Goldstein & West, 1999) and the statistical structure of caregivers' naturally-occurring social behavior is related to advances in early language (Pereira, Smith, & Yu, 2008). Thus, from the earliest stages, language is a socially situated process.

REFERENCES

Boysson-Bardies, B. (1993). Ontogeny of language-specific syllabic production. In B. Boysson-Bardies et al., (Eds.), *Developmental Neurocognition: Speech and face processing in the first year of life* (pp. 353-363). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Boysson-Bardies, B., & Vihman, M. M. (1991). Adaptation to language: Evidence from babbling and first words in four languages. *Language, 67*, 297-320.

Bromberg-Martin, E. S., & Hikosaka, O. (2009). Midbrain dopamine neurons signal preference for advance information about upcoming rewards. *Neuron, 63*, 119-126.

Delgado, M. R., Nystrom, L. E., Fisell, C., Noll, D. C., & Fiez, J.A. (2000). Tracking the hemodynamic responses to reward and punishment in the striatum. *J. Neurophysiol., 84*, 3072-3077.

Doupe, A. J., Kuhl, P. K. (1999). Birdsong and human speech: Common themes and mechanisms. *Annu. Rev. Neurosci., 22*, 567-631.

Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychol. Sci., 19*, 515-523.

Goldstein, M. H., & West, M. J. (1999). Consistent responses of human mothers to prelinguistic infants: The effect of prelinguistic repertoire size. *J. Comp. Psychol., 113*, 52-58.

Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proc. Natl. Acad. Sci. U.S.A., 100,* 8030-8035.

Goldstein, M. H., Schwade, J. A., & Bornstein, M. H. (2009). The value of vocalizing: Five-month-old infants associate their own noncry vocalizations with responses from caregivers. *Child Dev., 80*, 636-644.

Gómez, R. (2002). Variability and detection of invariant structure. *Psychol. Sci., 13,* 431-436.

Gómez, R. & Maye, J. (2005). The developmental trajectory of nonadjacent dependency learning. *Infancy, 7*, 183-206.

Kent, R. D. (1981). *Language Behavior and Early Childhood* (pp. 105-126). New York: Elsevier.

King, A. P., West, M. J., & Goldstein, M. H. (2005). Non-vocal shaping of avian song development: Parallels to human speech development. Ethology, *111*, 101-117.

Kuhl, P. K. (2007). Is speech learning "gated" by the social brain? *Dev. Sci., 10*, 110-120.

Kuhl, P. K., & Meltzoff, A. N. (1996). Infant vocalizations in response to speech: Vocal imitation and developmental change. *J. Acoust. Soc. Am., 100*, 2425-2438.

Legerstee, M. (1990). Infants use multimodal information to imitate speech sounds. *Infant Behav. Dev., 13*, 343-354.

Levitt, A., & Wang, Q. (1991). Evidence for language-specific rhythmic influences in the reduplicative babbling of French- and English-learning infants. *Lang. Speech, 34*, 235-249.

Marcus, G. F., Vijayan, S., Bandi Rao, S., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science, 283*, 77-80.

Meltzoff, A. N., Kuhl, P. K., Movellan, J., & Sejnowski, T. J. (2009). Foundations for a new science of learning. *Science, 325*, 284-288.

Movellan, J. R., & Watson, J. S. (2002). The development of gaze following as a Bayesian systems identification problem. In *Proc. 2nd Int. Conf. Dev. Learn.* ICDL02 (pp. 33-40). Washington, DC: IEEE Computer Society.

Oller, D. K. (2000). *The Emergence of the Speech Capacity*. Mahwah, NJ: Erlbaum.

Pereira, A. F., Smith, L. B., & Yu, C. (2008). Social coordination in toddler's word learning: Interacting systems of perception and action. *Connect. Sci., 20*, 73-89.

Rost, G. C., & McMurray, B. (2009). Speaker variability augments phonological processing in early word learning. *Dev. Sci., 12*, 339-349.

Saffran, J. R., & Thiessen, E. D. (2003). Pattern induction by infant language learners. *Dev. Psychol., 39*, 484-494.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*, 1926–1928.

Tenenbaum, J. B., & Griffiths, T. L. (2001). Generalization, similarity, and Bayesian inference. *Behav. Brain Sci., 24,* 629-640.

Tricomi, E., Delgado, M. R., McCandliss, B. D., McClelland, J. L., & Fiez, J. A. (2006). Performance feedback drives caudate activation in a phonological learning task. *J. Cognit. Neurosci., 18*, 1029-1043.

Vihman, M. M. (1993). The construction of a phonological systemIn B. Boysson-Bardies et al., (Eds.), *Developmental Neurocognition: Speech and Face Processing in the First Year of Life* (pp. 411-419). Dordrecht, The Netherlands: Kluwer Academic Publishers.

West, M. J., & King, A. P. (1988). Female visual displays affect the development of male song in the cowbird. *Nature, 334*, 244-246.

Whalen, D. H., Levitt, A. G., & Goldstein, L. M. (2007). VOT in the babbling of French- and English-learning infants. *J. Phon., 35,* 341-352.

CHAPTER 4

THE ROLE OF LOCOMOTION IN THE SOCIAL FACILITATION OF VOCAL DEVELOPMENT

Abstract

Compared to infants who cannot move independently, locomoting infants vocalize more and engage in more frequent social interaction (Gustafson, 1984). What mediates the relation between the development of locomotion and vocal communication? We hypothesized that caregivers are sensitive to their infants' ability to move independently, thus manipulations of infant locomotor status will change parental behavior. Parents of newly locomoting infants are motivated to be more vigilant and as a result engage in more contingent social interaction, providing their infants with greater vocal learning opportunities. This in turn should facilitate vocal development in locomoting infants. In Experiment 1, we assessed the prelinguistic repertoires of two groups of age-matched 9-mo-old infants who differed in their ability to locomote. We found that independently locomoting infants had greater proportions of advanced vocalizations in their prelinguistic repertoires. The total number of vocalizations across the two groups was not different. In Experiment 2, we directly manipulated the locomotor status of 11-month-old infants. When mobile, infants' repertoires had a larger proportion of canonical syllables than when they were immobile. Caregiver behavior was also sensitive to infant locomotor capabilities. Mothers increased the number of multi-modal initiations (involving both verbal and non-verbal components such as handing the infant an object while naming it) when

their infants were in the mobile condition. In summary, parents of locomoting infants initiated more interactions with them and their infants produced higher proportions of advanced vocalizations. We propose that increases in locomotor ability facilitate vocal development through changes in the dynamic of parent-infant interaction.

Introduction

The onset of self -produced locomotion is perhaps one of the most dramatic changes in the first year of an infant's life. After a lifetime of limitations, this new biomechanical advantage literally opens up a world of possibility for the infant. Early development encompasses multiple ongoing change processes that scale the body as well as the brain of the infant organism. Within the dynamical systems approach to the study of development as laid out by Thelen and Smith (1994), no solitary factor is responsible for the trajectory of development. Instead, multiple activity dependent developmental processes act in tandem to integrate perception, action and cognition (Thelen and Smith, 1994). In this view, development is an emergent property of the self-organization of multiple synchronous processes. Self-produced locomotion is one such activity dependent developmental process that itself catalyses developmental changes in multiple domains such as social communication, perception and affect (Piaget, 1954).

An average infant has travelled half the length of Manhattan in crawling steps by the time she is 10 months old (Adolph, 2005). The wealth of experience that the developing infant brain (and body) accumulates as a result of these journeys is immense. Once infants start to locomote, they are able to explore the limits of their own motor skills as well as the opportunities that the surfaces they encounter afford for learning (Gibson, 1988). The ability to explore objects in the environment from multiple vantage points (Adolph, 1997) probably goes hand in hand with the

ability to interact with many more objects that are now within reach. Recent evidence has indicated that increasing motor skill such as self-sitting and object manipulation guide the perceptual learning of objects as 3-dimensional forms. (Soska, Adolph and Johnson, 2010). Similarly, individual variation in infants' actions on objects predicts their ability to discern the boundaries between novel objects in a segmentation task (Needham, 2000). Thus the ability to explore objects makes for more robust knowledge of the environment. The onset of locomotion also enhances infants' memory formation skills (Rovee-Collier, 1996; Herbert, Gross, and Hayne, 2006; Gross, Hayne, Perkins and MacDonald, 2006), and earlier onset of locomotor (walking) skills has been associated with greater increases in infant autonomy (Biringen et al. 2008) as well as with greater cognitive advancement. The effect of independent locomotion on cognition is perhaps best demonstrated by the difference in infants' ability to succeed at multiple variations of the A-not-B spatial search task as a function of locomotor experience (Horobin and Acredolo 1986; Bertenthal, Campos and Barrett, 1984; Kermoian & Campos, 1988; Bertenthal and Campos, 1990; Tao and Dong, 1997). It has been suggested that changes in the ability to locomote facilitate infant ability to attend to the most relevant task features (Campos et al., 2000).

Infants who engage in self-produced exploration of the environment also find themselves distal from their caregivers (Gustafson, 1984) and this experience further facilitates social and cognitive development (Campos et al., 2000). Socially, infants display greater sociability and affectionate behavior after they start locomoting independently (Campos, Kermoian, & Zumbahlen, 1992). Infants with locomotor experience display increased ability to follow point and gaze gestures than age matched control infants and generally show greater distal attentiveness to affective communication from parents or events such as parental departure

(Campos, Kermoian and Zumbahlen, 1992; Campos et al. 1997

The dynamic of parent-child interaction also changes, with both parents and infants undergoing behavioral changes such as increased interactive play after the onset of independent locomotion (Campos et al., 2000). Infants display increased social referencing (Campos and Stenberg, 1981). Parents increase their use of specific affective contours in verbal communication as well as their displays of physical affection (Campos, Kermoian, & Zumbahlen, 1992). Recent data has also indicated that newly locomoting infants use social referencing while making decisions about motor actions, especially when faced by uncertainty in terms of locomotor capability (Tamis-LeMonda et al., 2008). Thus infants at this age are developing the ability to integrate information from motor and social domains. Correspondingly, their social environment (caregivers) also starts to change in response. For example, caregivers tailor their verbal communication, to match the infant's locomotor ability and age (Karasik et al., 2008).

It is not surprising that caregivers are aware of their infants' locomotor abilities and provide social information to the infants in accordance with this knowledge. Given parents' ability to predict intentional behaviors in infants such as coordinated reaching and grasping by the time infants are 9 months old (Reznick et al., 1988), independent locomotion is probably viewed as goal-directed behavior by caregivers. Parents tend to believe that infants will attempt locomotor tasks that lie outside their proficiency level such as attempt to crawl down an impossibly steep slope (Ishak et al., 2007). Thus it seems reasonable to assume that an increase in parental vigilance would accompany infant self-produced locomotion. Infants are much more likely to cause themselves accidental harm once they are able to move around on their own. Indeed, there is some comparative data from both altricial and precocial species to indicate that

changing locomotor ability and the proximity of the infant to the mother changes maternal vigilance. For instance, in the Chimpanzee, the level of maternal vigilance is higher when a dependent infant is separated from its mother than when s/he was in contact with the mother (Kutsukake, 2006). Similarly, increased maternal reactivity/anxiety to infant departures has been noted in three different species of macaques (Maestripieri, 1993). In a precocial species, the Alaskan Moose, females with active juveniles are more vigilant than those with inactive young (White and Berger, 2001). In humans, does increased vigilance on the part of the caregiver(s) alter the dynamic of the social interaction between infants and their caregivers? If so, does this altered social dynamic influence other ongoing developmental processes, such as vocal communication, that rely on social learning?

Arguably the most important socially motivated process that is ongoing at this time is vocal learning. However, the interaction of independent locomotion with communication and language has received little attention. One reason for this is that development of early speech and language were thought to proceed in the absence of social learning. Early work in the field of vocal development looked at maturational changes of the oropharyngeal tract during the first year of life and correlated these with the production of successively more advanced sounds (Kent, 1976, 1981, 1992). These approaches did not seek the socio-environmental factors that drove the development of vocal communication.

There is now a wealth of data that highlights the ubiquitous nature of social interaction in infant vocal learning processes. Social interactions are characterized by highly salient and structured components that attract and utilize infant attention to facilitate linguistic development. Caregivers use exaggerated infant-directed (ID) prosodic contours in their speech to infants (Fernald & Simon 1984; Fernald et al., 1989) and ID speech has been shown to effectively

engage infant attention (Kuhl, Coffet-Corina, Padden, & Dawson, 2005) and modulate infant affect (Fernald, Kermanschachi & Lees, 1984). ID speech has also been shown to aid word segmentation (Thiessen, Hill, & Saffran, 2005) and support phonetic category learning (Werker et al., 2007). Similarly, contingent responsiveness from caregivers facilitates developmental transitions to more advanced sounds along the developmental trajectory of prelinguistic infant vocalizations. Recent work has shown that infants' vocal behavior is sensitive to perturbations of the social environment. Infants use contingent social feedback in shaping their prelinguistic repertoires towards language-general acoustic features such as mature infraphonology (Oller, 2000; Goldstein, 2003). Similarly, the learning of language-specific features such as phonology [the way the sounds of a language are structured] is also mediated by social reinforcement (Goldstein & Schwade, 2008).

Indeed, turn-taking vocal interactions form a very early feature of parent-infant face-to-face interaction (Papousek, Papousek, & Bornstein, 1985) and infants respond to social contingency from a very young age, as well as show sensitivity to familiar contingency levels based on the responsiveness of their caregivers (Bigelow, 1998; Bigelow & Rochat, 2006). Speech from a live tutor is necessary for infants to learn phonemic contrasts in a foreign language (Kuhl, Tsao, & Liu, 2003). It has been suggested that live interaction is facilitative because social contingency aids learning by focusing infant attention on relevant features of stimulus information (Kuhl, 2007).

The development of locomotion and vocalization are both ongoing in tandem during the first year of an infant's life. How does vocal learning interact with ongoing changes in infant biomechanical abilities? Iverson and Thelen (1999) argued that early sensorimotor linkages between speech and motor systems lead to their later cognitive interdependence. Later work

building on this idea suggested that motoric behaviors such as gestures facilitate language learning (Iverson & Goldin-Meadow, 2005). With regard to locomotion, recent work has indicated that independent walking drives significant increases in infants' social interaction and vocal behavior (Clearfield, 2010). Our study extends these findings by examining vocal learning as a socially situated process, with contributions from both caregiver and infant. Rather than view increases in social interaction as a result of locomotor development, we believe that the effects are bidirectional – the onset of locomotion changes the social environment in ways that foster the development of communication. In our view, infants and caregivers mutually collaborate to construct developmental advances in communicative competence.

The present study investigates the effects of the ability to locomote on infants' social interaction and vocal learning. The onset of independent locomotion creates significant changes in the dynamic of the social interaction between infants and caregivers; locomoting infants engage in more complex social interactions with their caregivers that include a higher number of vocalizations and smiles (Gustafson, 1984). Social interactions in turn, shape vocal development (Goldstein et al., 2010, Kuhl, 2007). Do changes in the social environment that result from direct changes in locomotor capacity impact the development of infant vocal learning? Specifically, would increased maternal vigilance, caused by the onset of independent locomotion in the infant, lead to an increase in the amount of social interaction initiated by the mother, thus driving vocal change? We hypothesized that increased maternal vigilance in response to increases in infants' independent movement would cause changes in the caregiver-infant dynamic. As locomoting infants get more distal, caregivers should initiate more social interaction with them, and this in turn will facilitate vocal development.

Maternal initiation has been studied in the past as a form of stimulation with facilitative

effects on social and cognitive development. Amount of maternal stimulation at 12 months of age is related to infant habituation efficiency (Riksen-Walraven, 1978), hinting at some role for this process in cognitive development. Further, infants who show greatest competence while exploring have mothers that pay greatest attention to objects and events in the environment (Belsky, Goode and Most, 1980). Previous research has shown that infant cognitive development is correlated with mothers' use of stimulation that involves objects in the environment (Clarke-Stewart, 1973). Belsky, Goode and Most (1980) divided overall maternal stimulation into "physical" and "verbal" stimulation and found that mothers showed a linear increase in the amount of verbal stimulation they directed at their infants between the ages of 9 to 18 months (Belsky et al.,1980). Infants who received increased overall maternal stimulation at 12 months, showed more competent play 2 months after experiment completion (Belsky et al., 1980). It is interesting to note that 9-18 months is precisely the age period during which most infants learn to locomote. The ability to move through the environment enables infants to create many more opportunities for social interaction with others, such as maternal labeling of objects. Might there be a coupling between infant locomotion and the observed increase in maternal initiation at this age?

If caregivers are sensitive to their infants' ability to locomote, then manipulations of infant locomotor status will change parental behavior. Parents of newly locomoting infants should be more vigilant and as a result both engage in more contingent interaction, as well as initiate more interaction, providing their infants with greater opportunities for vocal learning. This in turn should facilitate vocal development in locomoting infants. As a first step, in Experiment 1, we investigated the relation between locomotor status and prelinguistic vocal development. To do this, we assessed the prelinguistic repertoires of two groups of 9-month-old

infants who differed in their ability to locomote. Experiment 2 controlled for the effects of motor maturation on the relation between locomotor and vocal behavior by manipulating opportunities to locomote independently in 11-12 month old infants. If changes in locomotor ability directly influence the structure of the social interaction between mother and infants, then infants should change their vocal repertoires as a function of locomotor opportunity. Furthermore, if the relationship between locomotor status and vocal development is mediated by changes in the dynamic of the parent-infant dyad, then parents should change their behavior when infants moved independently, with infant opportunity to locomote driving an increase in overall maternal stimulation.

Experiment 1

Methods

Participants

Thirty infants (mean age, 9 months 15 days; range 8 months 28 days to 10 months 10 days) and their mothers participated in the study. The infants were divided into two groups (n = 15) based on their locomotor ability. These data were randomly sub-sampled from the identical familiarization periods of two age matched subject cohorts. Each of these studies involved a 30-minute naturalistic interaction session between a parent-infant dyad on two consecutive days (Goldstein and Schwade, 2008; Goldstein, Syal, and Schwade, *in review*). The data for EXPERIMENT 1 were obtained from an initial10-minute free-play baseline period on day 2.

Apparatus

The study took place in a 12 X 18 ft playroom containing toys and posters of animals/cartoon characters on the walls. The large size of the playroom allowed infants to move

away from their mothers such that we could assay the effects of infant locomotion on maternal behavior. Dyads were recorded with three remotely controlled wall-mounted digital video cameras, routed to a DV tape recorder by a video switcher that allowed selection of the best camera angle. Each infant wore a pair of denim overalls concealing a wireless microphone and transmitter that enabled us to obtain high-quality recordings of the infants' vocalizations. Maternal vocalizations were recorded from wireless lapel microphone with a transmitter concealed in a pouch at the mothers' waist.

Procedure

Infant vocalizations were recorded during a 10-minute unstructured play session in the playroom. Infants and mothers were free to move around the room. Mothers were asked to play as they would at home.

Coding and analysis of infant vocalizations

Infant vocalizations were counted and scored for quality using the infraphonological coding scheme described by Oller (2000). This system describes prelinguistic vocal forms using both acoustic parameters such as formant transitions and qualitative descriptors such as phonetic categories. Two infraphonological parameters of primary interest are resonance and consonant vowel-articulation. The division of infant sound types according to their resonance includes "quasi- or fully-resonant vowels". Quasi-resonant vowels are nasalized and lack full vocalic status because air is pushed through nasal passages and a closed/slightly open vocal tract. Resonant vowels are normally phonated and produced with the full resonance of the upper vocal tract. Consonant-vowel articulation is measured by the time lag between the production of a consonant and a vowel sound in a syllable. A "marginal syllable" involves a slow transition between the consonant and the vowel (>200ms) producing an unclear sound. A "canonical

syllable" has a fast (<200ms) consonant-vowel transition and meets all the acoustic criteria for fully formed adult speech. We further collapsed across infraphonological categories to create 3 categories of vocal production as outlined in the table below.

Table 4.1: Advanced sound categories and descriptions

Infant sound categories	Description of categories
Fully Resonant	Syllables that had a fully-resonant nucleus
Complex	Syllables that contained a consonant-vowel articulation
Canonical	Syllables with a fast (<200ms) consonant-vowel articulation and were fully resonant. Canonical sounds meet all the criteria for adult speech.

We calculated three proportions corresponding to each of the above sound categories to characterize each infant's vocalization quality: (a) ratio of fully resonant vocalizations to total number of vocalizations, (b) ratio of vocalizations with CV structure to total number of vocalizations, and (c) ratio of canonical syllables to total number of CV-structured syllables.

Results

The total number of vocalizations did not differ across the two groups ($M_{non\text{-}loco} = 25.08$, SD = 14.0, $M_{loco} = 34$, SD = 16.5), $t(28) = 1.5965$, $p = .12$). A 2(locomotion) X 3(vocalization type) mixed ANOVA revealed a main effect of locomotion, $F(1,28) = 13.97$, $p = .001$ and a main effect of vocalization type, $F(2,56) = 10.67$, $p < .001$ (Figure 4.1). Significant main effects were followed up with Bonferroni-corrected planned comparisons of the effect of locomotion at each level of vocalization type (corrected alpha = .017). Locomoting infants produced significantly

greater proportions of resonant and canonical vocalizations than did nonlocomoting infants, $t_{resonance}(28) = 3.28, p = .003$; $t_{canonicity}(28) = 4.19, p < 0.001$. There was no significant effect of locomotion on vocalization complexity, $t_{complexity}(28) = 2.29, p = .03$.

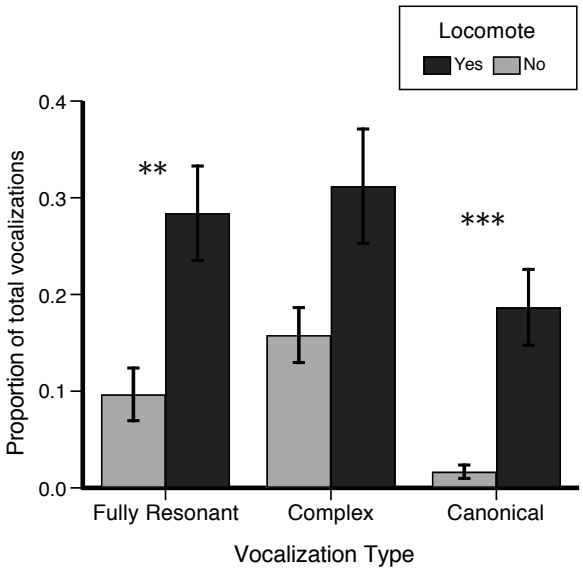

Figure 4.1: Locomoting infants produce significantly higher proportions (Mean $\pm$ 1 *SE*) of resonant and canonical utterances than age-matched non-locomoting counterparts. Asterisks indicate a significant difference between groups, $^{**}p < .01$, $^{***}p < .001$.

Discussion: Experiment 1

Experiment 1 examined connections between locomotor and vocal ability. We compared the vocal repertoires of two groups of infants who did not differ in terms of age or total number of vocalizations. We analyzed the effects of infants' locomotor status on the acoustic form of their prelinguistic vocalizations. Infants who were able to locomote produced higher proportions of fully-resonant and canonical vocalizations. While locomoting infants did not produce a higher proportion of more complex sounds than non-locomoting infants, our complexity category was the broadest measure of vocal quality, in that it contained the widest developmental range of

sound types. The measure included all sounds with consonant-vowel articulation, including both canonical (mature) and marginal (immature) syllables. The developmental breadth of the complexity category likely obscured differences between the locomotor groups.

With respect to canonical and resonant vocalizations, it is important to note that these infants did not produce more sounds than their age-matched non-locomotor counterparts, rather they produced a higher proportion of *more advanced* sounds. Thus the ability to locomote is related to the acoustic form of prelinguistic vocalizations. This effect is unlikely to be due to maturational differences because the ages of the infants in our locomoting and non-locomoting groups are not significantly different. However, this does not rule out the possibility that maturation of the articulators is related to motor maturation in general. To address this, in Experiment 2, we directly manipulated the locomotor status of a single group of infants.

Experiment 2

Participants

Sixteen infants, 8 girls and 8 boys (mean age, 11 months 13 days; range, 10 months 28 days to 12 months 2 days) and their mothers participated in the study. All infants were able to crawl independently, 1 infant was also able to walk. The data from an additional 6 infants could not be used in the as they became upset during the procedure.

Apparatus

The same apparatus as Experiment 1 were used. An additional upright activity center was added to the experimental playroom. The activity center was placed along the center of one wall of the room and was not moveable.

Procedure

Infants engaged in a 30-minute naturalistic interaction session with their mothers, divided into three 10-minute periods: Baseline, Immobile and Mobile. The order of the Immobile and Mobile periods were counterbalanced across infants. During the Baseline period, mothers were asked to play as they would at home. The purpose of the play session was to familiarize the infant with the playroom and toys. During the session, the infant was allowed to explore the activity center but was not placed in it. During the Immobile period, the infant was placed in the upright activity center. During the Mobile period, the infant was allowed to freely move around the room. The mothers were asked to remain seated on a cushion next to the activity center and fill out a background questionnaire and the MacArthur Communicative Development Inventory (CDI; Fenson et al., 1994) during both the immobile and mobile periods. This was done to simulate a more naturalistic setting wherein mothers are not constantly engaged in an exclusive interaction with their infants. The mothers were told that they were free to interact with their infants but that they should also focus on getting the questionnaires done.

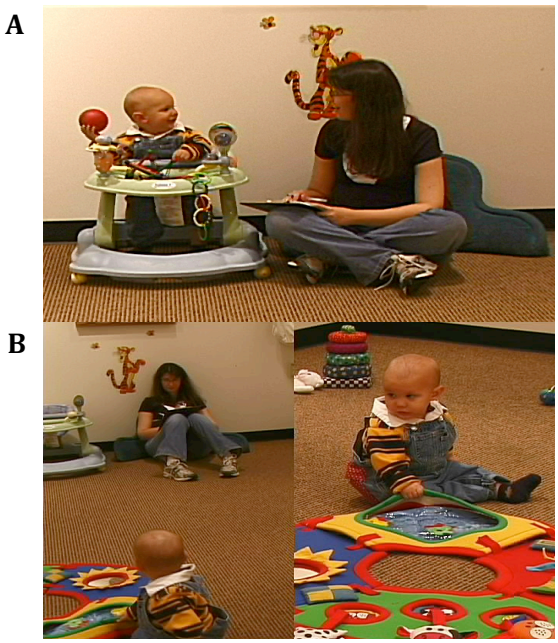

Figure 4.2 Experimental setup, A) Immobile condition, B) Mobile condition.

Data coding and analysis

Infant vocal behavior: Infant vocal behavior was assessed using the same infraphonological coding scheme as in Experiment 1. In addition to the first author, 4 coders were trained in infraphonological coding and coded 43% of the data. Average pairwise intercoder reliability was 90% (range: 88-92) Additionally, interrater reliability analysis using Cohen's Kappa was performed to determine consistency among raters, average Kappa = 0.86 (range: .837 - .888) , ps < .001.

Locomotion-associated vocalizations: To assess the temporal contingencies between locomotion and vocalizing in the mobile period, all the vocalizations that occurred during active locomotion, or in the 5 seconds immediately before or after a locomotor event, were coded as locomotion-associated vocalizations. The purpose of this coding scheme was to assess the relationship between vocalizing and locomoting and whether either reliably followed the other. A 5 second period was selected because contingent responses from caregivers that facilitate social learning generally occur within 5 seconds of infant vocalizations (Bornstein et al., 1992). Therefore a latency of 5 seconds between two events (in this case locomoting and vocalizing) is one that infants can detect and learn from.

Maternal behavior: Maternal behavior was divided into responses and initiations. Responses occurred within 3 seconds of an infant vocalization whereas the timing of initiations was not related to infant vocalizations. Both categories of maternal behavior were further subdivided into verbal, non-verbal and multimodal maternal behaviors. Verbal behaviors were coded when a mother spoke to her infant. Non-verbal maternal behavior was coded when the mother engaged in motor behavior such turning to look at her infant without saying anything. Multimodal

behaviors involved both vocal and non-vocal components such as handing her infant a toy while naming it. For example, if the mother spontaneously addressed her infant while pointing at a Winnie the Pooh poster and said "That's Pooh!", it was coded as a multimodal maternal initiation. If the mother responded to an infant babble directed at Winnie the Pooh and pointed and said "Is that Pooh?", it was coded as a multimodal maternal response.

Locomotion-associated vocalizing: As with infant behavior, maternal behavior directed at infants was coded during the 5 second preceding, the 5 seconds following and during infant locomotion in the mobile period. Average pairwise intercoder reliability for two additional trained coders was 93% (range: 91-95) with the first author. Additionally, inter-rater reliability analysis using Cohen's Kappa was performed to determine consistency among raters, average Kappa = 0.89, (range: .868 - .91) ps < .001.

Results

Maternal Behavior - Initiation: A 2 (locomotion: immobile, mobile) X 3 (initiation category: verbal, nonverbal, multimodal) repeated-measures ANOVA on number of maternal initiations revealed a main effect of initiation category, $F(2,30) = 14.11$, $p < .001$ and a significant interaction between infant locomotion condition and initiation category, $F(2,30) = 3.175$, $p = .056$ (Figure 4.4A). To decompose the interaction, we conducted 3 tests of simple main effects (Gravetter and Wallnau, 2009). For multimodal initiations, there was a significant effect of locomotion condition, $F(1, 15) = 4.99$, $p = .04$. Mothers directed more multimodal initiations at their infants when they were mobile than when they were not allowed to move around. For verbal initiations, there was a marginal effect of locomotion condition, $F(1, 15) = 3.829$, $p = .069$. Mothers tended to make fewer verbal initiations in the mobile period. For nonverbal initiations, there was no significant effect of locomotion condition $F(1, 15) = .155$, p

= . 699.

Maternal Behavior - Response: A 2(locomotion: immobile, mobile) X 3(response category: verbal, nonverbal, multimodal) repeated measures ANOVA on number of maternal responses revealed a main effect of response category, $F(2,30) = 9.58, p = .001$ and a significant interaction between infant locomotion condition and maternal response category, $F(2,30) = 3.91, \ p = .031$ (Figure 4.4B). We decomposed the interaction with tests of simple main effects. For non-verbal responses, there was a significant main effect of locomotion condition on number of maternal responses, $F(1, 15) = 5.48, p = .03$. Mothers directed more responses to their infants when they were mobile than when they were not allowed to move around. There were no significant main effects of locomotion condition for mothers' verbal and multimodal response categories, $F_{verbal}(1, 15) = 3.25, p = .09$, $F_{multimodal}(1, 15) = .081, p = .78$.

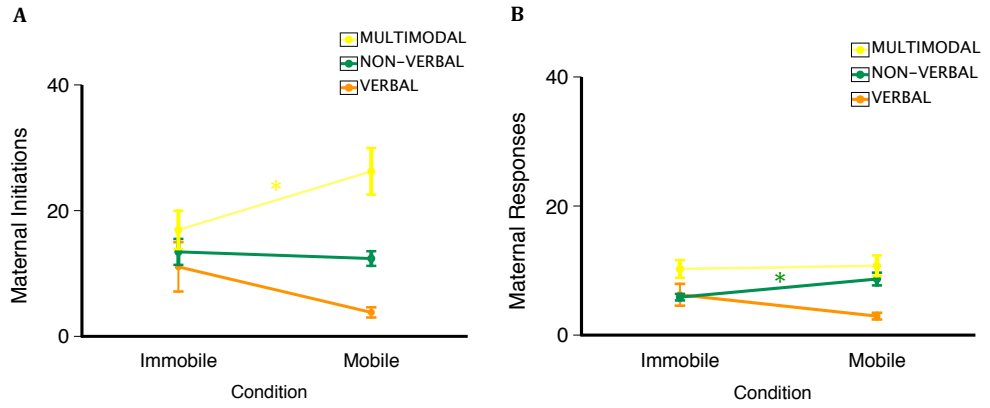

Figure 4.3: Specific changes in maternal behaviors (initiations and responses) across locomotion conditions. A) Mean number of maternal initiations, in each of the three initiation categories (verbal, nonverbal and multimodal) across the two test periods. B) Mean number of maternal responses, in each of the three response categories (verbal, nonverbal and multimodal) across the two test periods. Error bars represent $\pm$ 1*SE*. *$p < .05$

Mobile Period:

Maternal Initiations: Within the mobile period, infants were not moving all the time; locomotor events were interspersed with stationary bouts. We tabulated infant and maternal behaviors in the locomotion-associated period that consisted of the locomotion event as well as the 5 seconds immediately before and after the locomotion event. Mothers made a higher number of initiations in the locomotion-associated period ($M = 29.31$, $SD = 17.62$) than when their infants were not actively locomoting ($M = 10.88$, $SD = 12.62$), $t(15) = 2.949$, $p < .01$ (Figure 4.5A). To further assess differences in the number of initiations mothers made during the locomotion-associated period, we conducted a one-way repeated-measures ANOVA (locomotion-associated period: 5 sec before locomotion event, during locomotion event, 5 sec after locomotion event) on the proportion of maternal initiations from the mobile condition that occurred during each of the three locomotion-associated periods. There was a main effect of locomotion-associated period on maternal initiations, $F(2, 30) = 6.92$, $p = .003$ (Figure 4.5B). The proportion of initiations mothers made during locomotion events was significantly higher than in the 5 second periods before and after locomotion, Tukey's *HSD*, $ps < .01$. The proportion of maternal initiations in the 5 seconds before or after locomotion periods did not significantly differ from each other. To assess the relationship between locomotion and specific categories of maternal initiations, we conducted planned comparisons for each initiation category using separate one-way ANOVAs with repeated measures on locomotion period. We found a significant main effect of locomotion on non-verbal initiations in the locomotion-associated period $F(2, 30) = 7.612$, $p = .002$ (Figure 4.5C). The proportion of nonverbal initiations while the infant was actively locomoting exceeded those in the 5-second time windows before and after locomotor events, Tukey's *HSD*, $ps < .01$.

Maternal Responses: Similar to the pattern of initiations, mothers made a higher number of overall responses in the locomotion-associated period ($M = 19.38$, $SD = 3.88$, $SD = 11.95$) than when infants were stationary ($M = 3.88$, $SD = 6.238$), $t(15) = 4.096$, $p < .001$ (Figure 4.5D). A one-way repeated-measures ANOVA (locomotion-associated period: before, during, after locomotion event) on the proportion of maternal responses revealed a main effect of locomotion-associated period, $F(2,30) = 3.826$, $p = .033$ (Figure 4.5E). Post-hoc tests indicated that the proportion of responses mothers made during locomotion was significantly higher than (< 5 sec) after locomotion but not significantly different from (< 5 sec) before locomotion, Tukey's *HSD* $ps < .05$. To assess the effect of locomotion-associated period on specific categories of maternal responses, we conducted separate planned comparisons for each response category using one-way ANOVAs on the proportion of maternal responses with repeated measures on locomotion-associated period. Again, we found a significant effect of locomotion-associated period only for non-verbal responses, $F(2, 30) = 5.075$, $p = .013$ (Figure 4.5F). Mothers' proportion of nonverbal responses while infants were actively locomoting exceeded those in the 5-second time window after locomotor events but not in the 5 seconds before locomotion, Tukey's *HSD*, $ps < .05$.

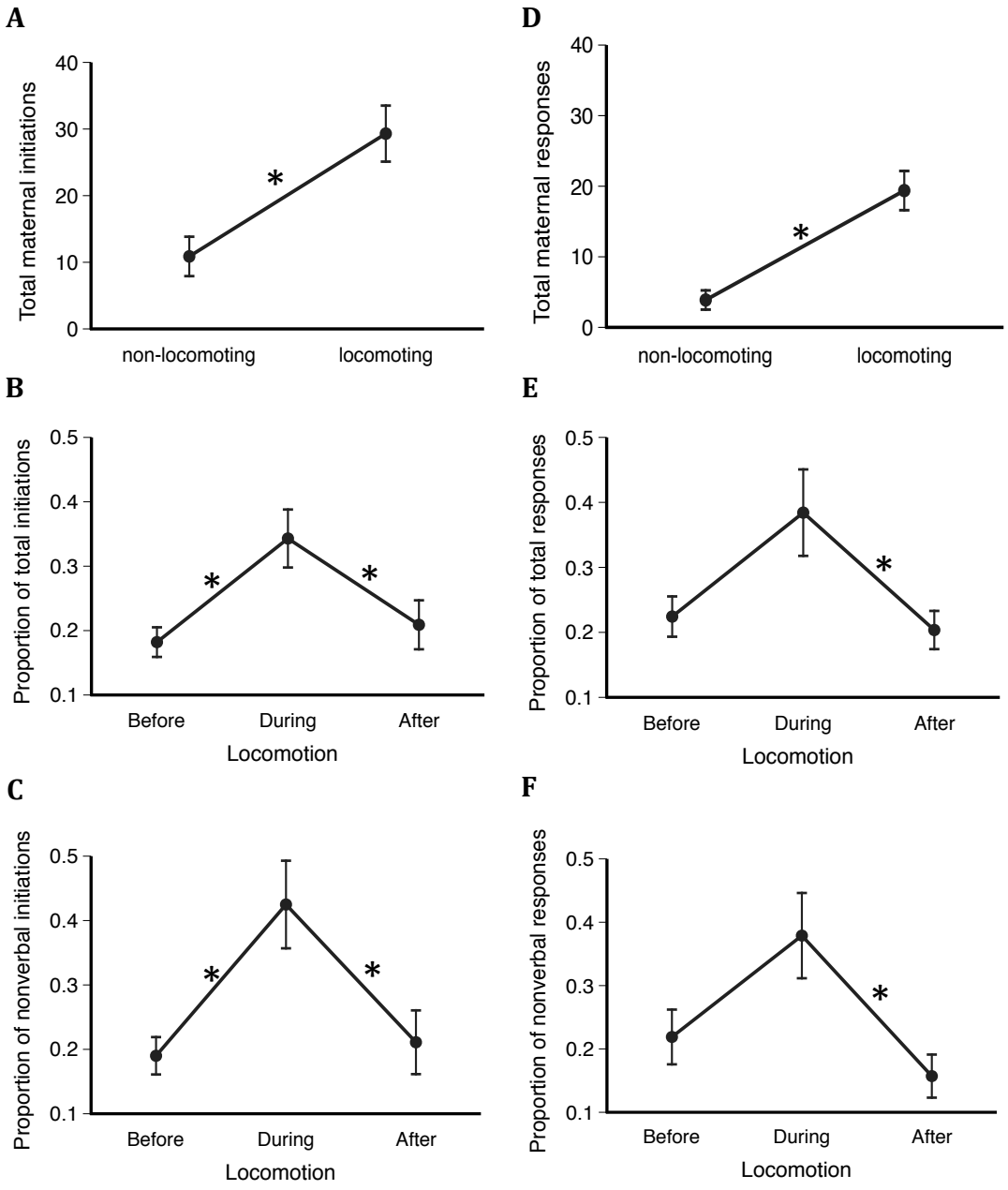

Figure 4.4 Changes in maternal behavior in response to infant locomotion. A) Difference in overall number of maternal initiations between stationary and locomotion associated periods. B) Changes in the mean proportion of total initiations within each locomotion-associated period, C) Changes in the mean proportion of nonverbal initiations within each locomotion-associated period, D) Difference in overall number of maternal responses between stationary and locomotion associated periods. E) Changes in the mean proportion of total responses within each locomotion-associated period, D) Changes in the mean proportion of non-verbal responses within each locomotion-associated period. Error bars represent ± 1*SE*. **p* < .05

Infant Behavior: A 2(locomotion: immobile, mobile) X 3(vocalization type: fully-resonant, complex, canonical) repeated-measures ANOVA on proportion of vocalizations revealed a main effect of locomotion, $F(1,15) = 13.804$, $p = .002$ and a main effect of vocalization type, $F(2,30) = 18.712$, $p < .001$. To investigate whether these main effects were due to specific forms of vocalizing, we performed planned comparisons on locomotion condition within each level of vocalization type (Bonferroni-corrected alpha = 0.0167). When mobile, infants produced significantly higher proportions of vocalizations that met the criteria for Resonance, $t(15) = -3.101$, $p = .007$, and Canonicity, $t(15) = -3.298$, $p = .005$. There was no effect of locomotion condition on vocal complexity, $t(15) = -2.653$, $p = .019$; (Figure 4.6)

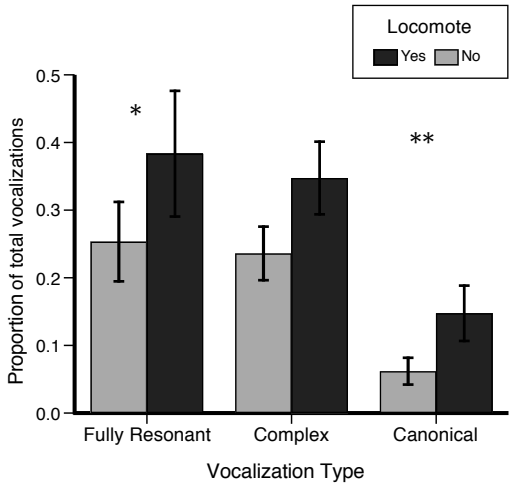

Figure 4.5 Infants produced significantly higher proportions (Mean $\pm$ 1 *SE*) of resonant and canonical utterances in the mobile period than in the immobile period. *$p < .05$, **$p < .01$.

Mobile Period

Infants made a higher proportion of vocalizations during the locomotion-associated period ($M = 0.67$, $SD = 0.29$), than when they were not locomoting ($M = 0.33$, $SD = 0.28$), $t(15) = 2.243$, $p = .04$ (Figure 4.7A) Within the locomotion-associated period, a one-way repeated-

measures ANOVA (locomotion-associated period: before, during, after locomotion events) on proportion of vocalizations yielded a significant main effect of locomotion $F(2, 30) = 12.41$, $p < .001$. Post-hoc tests revealed that revealed that the proportions of infant vocalizations made during locomotion were significantly higher than before or after locomotion, Tukey's HSD, ps < .05 (Figure 4.7B).

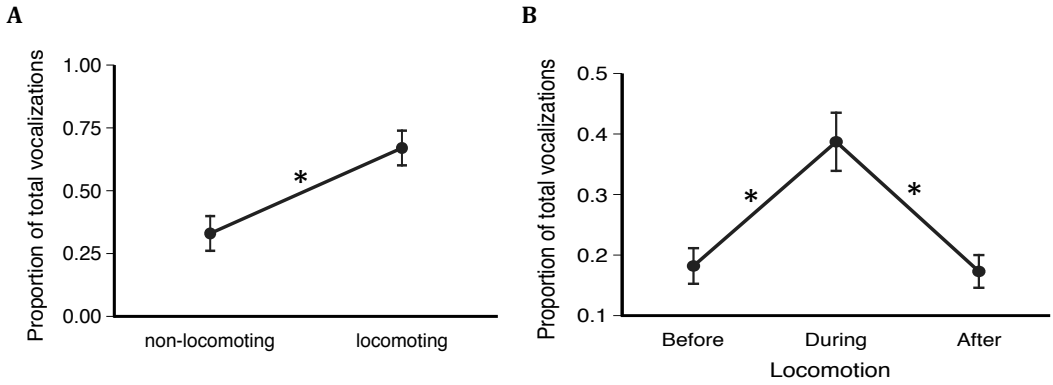

Figure 4.6: Changes in proportion of infant vocal behavior across locomotion periods. A) Differences in mean proportion of total vocalizations across locomotion conditions, B) Changes in the mean proportion of total vocalizations within the locomotion-associated period. Error bars represent ± 1*SE*. * $p < .05$

Mobile Period - Learning effect

We also analyzed learning within the mobile period by comparing the proportion of advanced vocalizations during early versus late time bins within the mobile period. We divided infant vocalizations into 4 time bins within the mobile period, spanning 2.5 minutes each (total mobile period = 10 minutes). We then calculated proportions of "resonant", "complex" and "canonical" sounds that occurred within these time bins. We compared the proportion of advanced sounds between early and late time bins, which were:

1. Time Bin 1 (0-2.5 minutes) versus Time Bin 4 (7.5-10 minutes)

2. Time Bin 1_2 (0-5 minutes) versus Time Bin 3_4 (5-10 minutes)

A repeated measures 3(Voc_Type) X 2 (Time_Bin) ANOVA did not reveal a significant any significant main effects or interactions. Infant vocalizations are not significantly different across early and late time bins.

Discussion: Experiment 2

Experiment 2 consisted of a real-time within-subjects manipulation of locomotion opportunities to control for any maturational effects that may mediate the connection between locomotor and vocal ability. Infants were placed in both immobile and mobile contexts in order to assess the effects of the ability to locomote on both maternal and infant behavior. Our hypothesis was that when infants were mobile, increased vigilance would cause mothers to initiate more social interaction with them and this would in turn facilitate the production of more advanced sound forms by the infants.

Mothers directed a higher proportion of initiatory and responsive behavior at their infants when they were in the mobile condition. Mothers increased both their overall number of initiations as well as multimodal initiations. There was a marginal decrease in the proportion of verbal-only initiations from the mothers during the mobile period. In terms of responses, mothers increased non-verbal responses i.e. looks towards their infants in the mobile condition, supporting the idea that mothers were more vigilant when their infants are mobile.

We then assessed changes in both maternal and infant behavior in the locomotion-associated period during which infants were actively engaged in moving from one place to another within the experimental playroom. We found that mothers directed more overall initiations and responses at their infants during the locomotion associated periods than when the infants were stationary. Specifically, mothers made the highest proportion of initiations *during* locomotion, as compared to before or after locomotion during the locomotion-associated period.

Non-verbal maternal initiations showed the greatest increase. Maternal responses during the locomotion-associated period followed the same general pattern as initiations. Thus both nonverbal initiations and nonverbal responses were significantly higher during active locomotion, lending strong support to our vigilance hypothesis.

In terms of infant vocal quality, the results of Experiment 2 replicated and extended those of Experiment 1; infants produced higher proportions of resonant and canonical vocalizations in the mobile condition. In our assessment of the temporal relationship between locomotion and vocalization in infants, we found that a greater proportion of infant utterances occurred in the locomotion-associated period. Within the locomotion-associated period, infants made the highest proportion of vocalizations during locomotion as compared to before and after locomotion.

General Discussion

Our experiments indicate that the changing dynamic of parent-infant interaction upon the onset of independent locomotion mediates the effect of locomotor ability on vocal ability. In Experiment 1, we found that infants who were able to crawl produced more advanced vocal forms than age-matched controls who were not able to locomote independently. It has been suggested that individual differences in the timing of developmental transitions, such as the onset of independent locomotion, can affect wide scale changes in infant behavior (Biringen, Emde, Campos & Appelbaum, 1995). In order to rule out these effects, we decided to assay the behavior of the same infants in locomotor and non-locomotor contexts in Experiment 2.

We found that mothers increased their total number of infant-directed behaviors (both responses and initiations) during the mobile condition. In terms of responses, they showed an increase in nonverbal responses to their infants in the mobile condition. Specifically, mothers

were most responsive while their infants were moving independently. Further, the highest proportion of initiations occurred during infant locomotion. Since non-verbal maternal behaviors were comprised of looks to the infant, all of these data support the idea that increased maternal vigilance when the infant is locomoting brings about specific changes in maternal behavior. The role of eye gaze in vigilance-related arousal is well known and is used as an index for anxiety (Bradley at al., 2000; Mogg & Bradley, 2002). Thus increased looks directed at the infant when s/he was mobile provide credence to the idea that mothers of locomoting infants are more vigilant. Vigilance and attention are intrinsically linked (Lim and Dinges, 2008) and increased attending to the infants causes increased maternal behavior directed at the infant, thus facilitating vocal development.

Additionally, mothers made a higher number of multimodal initiation when their infants were mobile. Multimodal initiations commonly involved labeling something in the room (toy or picture). Mothers' increase of this behavior towards locomoting infants ties in well with the idea that caregivers are sensitive to changes in infant context. Locomoting infants interact with more objects and are more social (Gustafson, 1984). In turn, mothers seem to initiate more complex or multi-modal interactions and contain both verbal and non-verbal (looking/pointing/touching) components. Mothers also showed a marginal decrease in the number of verbal initiations they made in the mobile condition. While this may seem surprising, we think it derives from the increased need to attract infant attention using multimodal cues. When the infant is immobile, s/he is most likely to pay attention to the objects directly in front of him/her but this changes when the infant is able to independently explore the environment. Previous work has shown that locomoting infants direct increased attention at distal objects or events (Gustafson, 1984; Campos et al., 2000). Gustafson noted that infants approached and explored previously

inaccessible objects when allowed to move independently. Locomoting infants also look significantly more at the parent and at distant toys as well as non-toy features of the room such as posters on the wall (Gustafson, 1984).

Locomoting infants show an increased number of directed behaviors such as vocalizing and smiling at adults (Gustafson 1984). These changes in infant perception and action are met by changes in maternal behavior in ways that help to organize infant attention and facilitate infant learning. Increased interactive play and social referencing at this time (Campos, et al., 2000, 1997; Tamis-LeMonda, 2008) serves to redirect infant attention to both proximal and distal learning opportunities in the environment. An early study documented that 56% of maternal vocalizations directed at 12 mo old infants were in response to infant exploratory behaviors with mothers either questioning or labeling whatever the infant was looking at along with frequent maternal labeling of objects upon the infant's first encounter with them (West and Rheingold, 1978). Additionally, mothers spontaneously change their behavior to be more repetitive, interactive and reinforcing when demonstrating new objects to 11-13 month-old infants (Brand et al., 2002)

Thus once the infants are able to explore new objects in the world independently, mothers needs to combine speech with action (multimodal initiation) in order to channel infant attention to new learning opportunities. This ties in well with the literature on the role of intersensory redundancy in early infant word-object mapping. Mothers typically provide infants with spatial and temporal synchrony when presenting object labels and this helps object-referent matching. (Gogate and Bahrick, 1998: Gogate, Bahrick and Watson, 2000; Gogate, Betancourt, Bolzani, and Watson, 2000). One way in which mothers can aid invariance detection is by pointing to or handing infants' objects while labeling them and this is evidenced in their increase of multimodal

initiations at the cost of verbal ones. It is conceivable that mothers label the same objects in multiple contexts (such as rooms in the houses) once the infant starts locomoting and that this further aids infant vocal learning.

It has recently been proposed that increased interaction with objects "sets a context for attributing meaning to those objects via action" (Iverson, 2010). Is this aided by independently locomoting infants' ability to interact with more objects? Karasik, Tamis-LeMonda and Adolph (under review) have noted that the onset of walking is associated with a change in the way that infants interact with objects in the home environment – walking infants are more likely to bring an object to their mother than crawling infants. Our study did not permit assaying the differences between object interaction (number of objects, handling, object directed vocalizations etc.) during the immobile and the mobile conditions because infants had access to many toys in the mobile condition but only a limited subset of those in the immobile condition. A comparison of object-directed behavior between infants in mobile and immobile conditions with identical objects might shed light on the interaction of vocal and locomotor development as a result of changes in infant ability to interface with objects in the environment.

While we expected to see increased overall maternal responsiveness to infant vocalizations in the mobile condition, our data did not indicate that this was the case. Maternal increases in responsiveness were limited to non-verbal responses that involved looks to the infant. Other work has documented infant ability to learn more advanced forms of vocalizing from silent caregiver responses (Goldstein et al., 2003) and one possibility is that the current findings may derive from similar mechanisms. A second possibility is that the relative proportions of mother-infant interaction that are formed by responses and initiations change at the time of independent locomotion onset in infant development. Mothers might increase the

number of initiations in order to direct infant attention towards learning opportunities. A third possibility is that mothers become more attuned to newly locomoting infants' exploratory actions on the environment and shift some proportion of their responses to infant actions as opposed to infant vocalizations.

Infants in Experiment 2 produced higher proportions of resonant and canonical (advanced) vocal forms when they were mobile, akin to the findings of Experiment 1. Since Experiment 2 used a within-subjects manipulation, it is highly unlikely that simple maturational effects (that could have caused some infants to be more advanced in both motoric and linguistic domains) drove the results of Experiment 1. The changes in infant vocal behavior when mobile were intimately tied to changes in maternal behavior (outlined above). Previous work has shown that specific changes in maternal behavior guide infant vocal development. Contingent responses from caregivers facilitate the production of more advanced sounds from prelinguistic infants (Goldstein et al., 2003; Goldstein and Schwade, 2008).

Within the mobile period, infants made a significantly higher number of vocalizations during the locomotion-associated period with a significantly higher proportion of vocalizations coinciding with actual locomotion. Typically, an infant would spot a distal object or the parent and then vocalize while moving towards this point of interest. Similar tight coupling between vocal and motor behavior has been noticed in the links between rhythmical movements of the arms and hands and reduplicated babbling in infants (Oller & Eilers, 1988; Thelen, 1996; Locke et al., 1995, Iverson & Thelen, 1999). Infants show a sharp increase in rhythmic arm shaking at the time of reduplicated babbling onset (Iverson, Hall, Nickel & Wozniak, 2007). It has been hypothesized that early sensorimotor linkages between cortical domains involved in vocalizing and motor movements of the hand lead to later cognitive interdependence between these

activities (Iverson & Thelen 1999).

Early thought about the interaction between motor and linguistic function was largely been dominated by the frame/content theory (MacNeilage) according to which modifications of the ventral and supplementary pre-motor cortices allowed for the evolution of "speech frames" from within the repertoire of mandibular cycles. However, this view does not account for reciprocal influences in motor and linguistic advancement. From a neuroanatomical point of view, it is likely the motor and language systems develop in tandem. Long distance cortico-cortical connections have been found between regions that subserve language and action such as the links between left inferior frontal (Broca's) and superior temporal (Wernicke's) and dorsal and ventral premotor cortices (Hickok & Poeppel, 2004). In addition the cerebellum, most commonly identified with movement processes, both projects to traditional language areas such as Broca's (Leiner et al., 1989, 1993) and is activated during word production or word association tasks (Peterson et al., 1989). Somatotopic mapping of the supplementary motor cortex indicates that areas that are involved in arm and hand movement lie adjacent to areas that regulate speech production (Fried et al., 1991).

Further support of the idea that motor and linguistic processes might be served by distributed interactive functional systems comes from findings of embodied language processing in adults. For example, reading an action word (such as "kick") activates pre-motor and motor regions that are involved in the actual execution of the action (Hauke, Johnsrude & Pulvermuller, 2004). The converse of this is also true, with activation in language-related regions (parts of Broca's area) during motor hand movement (Erhard et al., 1996) as well as thinking about motor hand movements (Krams et al., 1998). All of these argue for combinatorial processing of linguistic and motor processing in the brain that leads to the emergence of streamlined/seamless

interactions between these two domains.

Relatedly, adults find it easier to make sensibility judgments when task completion requires a physical action that is in the same direction as an action outlined in the stimulus sentence (Glenberg & Kaschak, 2002). These results lend support to the idea that language is embodied in the way that we perform the actions that are described in our words, and this has found evidence in infants' early word learning as well. Children's initial uses of verbs encode their own actions. Huttenlocher, Smiley and Charney, (1983) reported that in a sample of 16, 1 yr 10 mo to 3 yr 6 mo old children, 90 % of the total of 1066 utterances with verbs were produced when the child was participating in the action. Similarly, when 2.5 yr old children are asked to make category judgments about actions, they do so on the basis of their own actions on those objects (Smith, 2005).

Within the embodied cognition perspective, the mind is an emergent property of the interaction between the body and the environment it lives and moves in (Thelen, 2000). It follows then, that changes in the ability to move through the environment would impact infant learning of cognitive faculties such as language. The dynamical systems approach posits that infants parse regularities in the world from the cross-correlation of multimodal experiences i.e. the detection of structural or statistical patterns in experiences with the environment help infants learn about the environment. However, it seemingly ignores one very crucial part of the infants experience with the world – the caregiver. Of the multiple synchronous processes that self-organize for development to emerge, one process, is infant-caregiver interaction. It has recently been hypothesized that two criteria must be met for infants to learn new sources of structure. One is statistical regularity within a small time window (such as contingent and repetitive object-label assignment) and the other is behavioral significance such as social salience (socially reinforcing

label assignment from the parent) (Goldstein et al., 2010a). Social interaction increases *learnability* by salience-tagging what needs to be learnt. In this view, parental behavior aids embodiment by grounding early language in action. Parental presentation of links between sounds and words to their infants coincides with their offering matched repetitive demonstrations of an action that is denoted by the word. For example, saying "round and round and round and round", while demonstrating how a tape is wound around a tape roll (Yoshida and Smith, 2005). An approach to dynamical systems theory that incorporates social context (i.e. changing maternal behaviors as a control parameter in infant development) is a more accurate representation of the environment that the infant learns from and in.

Our attempt in the present experiments was to study the infant as part of a social system. We reasoned that changes in the social environment would drive changes in the infant. Maternal behavior is sensitive to infant locomotion and the onset of independent locomotion in the infant changes maternal motivations. Increased maternal vigilance leads to increased social interaction between the mother and the infant. This change in caregiver behavior in turn creates more opportunities to learn which infants capitalize on, thus propelling themselves forward on the trajectory of vocal development.

REFERENCES

Adolph, K. E. (1997). Learning in the development of infant locomotion. *Monographs of the Society for Research in Child Development, 62*(3, Serial No. 251).

Adolph, K. E. (2005). Learning to learn in the development of action. In J. Lockman & J. Reiser (Eds.), *Action as an organizer of perception and cognition during learning and development: The 32nd Minnesota Symposium on Child Development* (pp. 91-122). Hillsdale, NJ: Lawrence Erlbaum Associates.

Adolph, K. E. (2008). Learning to move. *Current Directions in Psychological Science, 17*, 213-218.

Belsky, J., Goode, M. K., & Most, R. K. (1980). Maternal stimulation and infant exploratory competence: Cross-sectional, correlational, and experimental analyses. *Child Development, 51*, 1163–1178.

Bertenthal, B. I., & Campos, J. J. (1990). A systems approach to the organizing effects of self-produced locomotion during infancy. In C. Rovee Collier & L. P. Lipsitt (Eds.), *Advances in infancy research* (Vol. 6, pp. 1–60). Norwood, NJ: Ablex.

Bertenthal, B. I., Campos, J. J., & Barrett, K. C. (1984). Self-produced locomotion: An organizer of emotional, cognitive, and social development in infancy. In R. Emde & R.Harmon (Eds.), *Continuities and discontinuities in development* (pp. 175–210). New York: Plenum.

Bigelow, A. E. (1998). Infants' sensitivity to familiar imperfect contingencies in social interaction. *Infant Behavior and Development*, 21, 149–162.

Bigelow, A. & Rochat, P. (2006). Two-month-old Infants' Sensitivity to Social Contingency in Mother-Infant and Stanger-Infant Interaction. *Infancy, 9*(3), 313-325.

Bornstein, M. H., Tamis-LeMonda, C. S., Tal, J., Ludemann, P., Toda, S., Rahn, C. W., Pecheux, M., Azuma, H., & Vardi, D. (1992). Maternal responsiveness to infants in three societies: America, France, and Japan. *Child Development, 63*, 808–821.

Brand, R. J., Baldwin. D. A., & Ashburn, L. (2002). Evidence for "Motionese": Mothers modify their infant-directed actions. *Developmental Science*, 5, 72-83.

Burger, J., & Gochfeld, M. (1994). Vigilance in African Mammals: Differences Among Mothers, Other Females, and Males. *Behaviour, 131*(3-4), 153-169.

Campos, J. J., & Stenberg, C. (1981). Perception, appraisal, and emotion: The onset of social referencing. In M. Lamb & L. Sherrod (Eds.), *Infant social cognition: Empirical and theoretical considerations* (pp. 274-313). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Campos, J. J., Kermoian, R., & Zumbahlen, M. R. (1992a). Socioemotional transformations in the family system following infant crawling onset. In N. Eisenberg & R. A. Fabes (Eds.), Emotion and its regulation in early development (pp. 25-40). San Francisco: Jossey-Bass

Campos, J. J., Bertenthal, B. I., & Kermoian, R. (1992b). Early experience and emotional development: The emergence of wariness of heights. *Psychological Science, 3*, 61-64.

Campos, J. J., Kermoian, R., Witherington, D., Chen, H., & Dong, Q. (1997). Activity, attention and developmental transitions in infancy. In P. J. Lang & R. F. Simons (Eds.), *Attention and orienting: Sensory and motivational processes* (pp. 393– 415). Mahwah, NJ: Erlbaum.

Campos, J. J., Anderson, D. I., Barbu-Roth, M. A., Hubbard, E. M., Hertenstein, M. J., & Witherington, D. (2000). Travel Broadens the Mind. *Infancy*, *1*(2), 149-219.

Clarke-Stewart, K.A. (1973). Interactions between mothers and their young children: Characteristics and consequences. *Monographs of the Society for Research in Child Development, 38* (6-7, Serial No. 153).

Erhard, P., Kato, T., Strupp, J. P., Andersen, P., Adriany, G., Strick, P. L., & Ugurbill, K. (1996). Functional mapping of motor in and near Broca's area, *Neuroimage, 3*, S367.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59*(5, Serial No. 242), 1–173.

Fernald, A., & Simon, T. (1984). Expanded intonation contours in mother's speech to newborns. *Dev. Psychol., 20*, 106-112.

Fernald, A., Kermanschachi, N., & Lees, D. (1984). The rhythms and sounds of soothing: Maternal vestibular, tactile, and auditory stimulation and infant state. *Biennial*

International Conference on Infant Studies.

Fernald, A., Taeschner, T., Dun, J., Papousek, M., de Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *J. Child Lang., 16*, 477–501.

Fried, I., Katz, A., McCarthy, G., Sass, K. J., Williamson, P., Spencer, S. S., & Spencer, D. D. (1991). Functional organization of human supplementary motor cortex studied by electrical stimulation. *Journal of Neuroscience*, *11*, 3656-3666.

Gibson, E. J. (1988). Exploratory behavior in the development of perceiving, acting, and the acquisition of knowledge. *Annual Review of Psychology, 39*, 1-41.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review, 9*(3), 558-565.

Gogate, L. J. & Bahrick, L. E. (1998). Intersensory redundancy facilitates learning of arbitrary relations between vowel sounds and objects in 7-month-old infants. *Journal of Experimental Child Psychology*, *69* (2), 133-149.

Gogate, L. J., Bahrick, L. E., & Watson, J. D. (2000). A study of multimodal motherese: The role of temporal synchrony between verbal labels and gestures. *Child Development*, *71*(4), 876-892.

Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proc. Nat. Acad. Sci., 100*(13), 8030-8035.

Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychol. Sci., 19*(5), 515-523.

Goldstein, M. H., Waterfall, H., Lotem, A., Halpern, J., Schwade, J. A., Onnis, L., & Edelman, S. (2010a). General cognitive principles for learning structure in time and space. *Trends in Cognitive Sciences*, 14 (6), 249 - 258.

Goldstein, M. H., Schwade, J. A., Briesch, J., & Syal, S. (2010b). Learning while babbling: Prelinguistic object-directed vocalizations signal a readiness to learn. *Infancy*, 15(4), 362-391.

Gravetter, F. J., & Wallnau, L.B. (2009). *Statistics for the behavioral sciences, 8th edition*. Belmont, CA: Wadsworth, Cengage Learning.

Gross, J., Hayne, H., Perkins, N., & McDonald, B. (2006) "Amount of Crawling and Walking Experience has Varying Effects on Cognitive Development During Infancy" Paper presented at the annual meeting of the *XVth Biennial International Conference on Infant Studies*, Westin Miyako, Kyoto, Japan.

Gustafson, G. E. (1984). Effects of the ability to locomote on infants' social and exploratory behaviors: An experimental study. *Developmental Psychology*, *20*(3), 397-405.

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in the motor and premotor cortex. *Neuron, 41*, 301-307.

Herbert, J., Gross, J., & Hayne, H. (2006). Age-related changes in deferred imitation between 6 and 9 months of age. *Infant Behavior and Development, 29*, 136-139.

Hickok, G., & Poeppel, D. (2004) Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language, Cognition, Volume 92(1-2), 67-99

Horobin, K., & Acredolo, L. (1986). The role of attentiveness, mobility history and separation of hiding sites on Stage IV search behavior. *Journal of Experimental Child Psychology, 41,* 114–127.

Huttenlocher, J., Smiley, P., & Charney, R. (1983). The emergence of action categories in the child: Evidence from verb meanings. *Psychological Review, 90*, 72-93.

Ishak, S., Tamis-LeMonda, C. S., & Adolph, K. E. (2007). Ensuring Safety and Providing Challenge: Mothers' and Fathers' Expectations and Choices About Infant Locomotion. *Parenting: Science and Practice*, *7*(1), 57-68

Iverson, J. M., & Thelen, E. (1999). Hand, mouth, and brain: The dynamic emergence of speech and gesture. *Journal of Consciousness Studies, 6*, 19-40.

Iverson, J.M. & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science*, 16, 367-371.

Iverson, J. M., Hall, A. J., Nickel, L., & Wozniak, R. H. (2007). The relationship between onset

of reduplicated babble and laterality biases in infant rhythmic arm movements. *Brain and Language, 101*, 198-207.

Kent, R. D. (1976). Anatomical and neuromuscular maturation of the speech mechanism: Evidence from acoustic studies. *J. Speech Lang.Hear. Res. 19*, 421– 447.

Kent, R. D. (1981). Articulatory-acoustic perspectives on speech development. In R. Stark (Ed.), *Language Behavior in Infancy and Early Childhood* (pp. 105-239). Amsterdam: Elsevier North-Holland.

Kent, R. D. (1992). Phonological development as biology and behavior.

In R.S. Chapman (Ed.), *Processes in Language Acquisition and Disorders* (pp. 67-85). St. Louis, MO: Mosby-Year Book.

Kermoian, R., & Campos, J. J. (1988). Locomotor experience: A facilitator of spatial-cognitive development. *Child Development, 59,* 908–917.

Krams, M., Rushworth, M. F. S., Deiber, M. P., Frackowiak, R. S. J., & Passingham, R. E. (1998). The preparation, execution and suppression of copied movements in the human brain, *Exp. Brain Res. 120*, 386–398.

Kuhl, P. K., Tsao, F. M. &, Liu, H. M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proc. Natl. Acad. Sci., 100*, 9096–9101.

Kuhl, P. K., Coffey-Corina, S., Padden, D., & Dawson, G. (2005). Links between social and linguistic processing of speech in preschool children with autism: Behavioral and electrophysiological evidence. *Dev. Sci., 8*, 1-12.

Kutsukake, N. (2006). The context and quality of social relationships affect vigilance behaviour in wild chimpanzees. *Ethology, 112*, 581-591.

Leiner, H. C., Leiner, A. L., & Dow, R. S. (1993). Cognitive and language functions of the human cerebellum. *Trends Neurosci, 16*, 444–7.

Maestripieri, D. (1993). Maternal anxiety in rhesus macaques (*Macaca mulatta*). I. Measurement of anxiety and identification of anxiety-eliciting situations. *Ethology, 95*, 19-31.

Needham, A. (2000). Improvements in object exploration skills may facilitate the development of object segregation in early infancy. *Journal of Cognition and Development, 1,* 131–156.

Oller, D. K. (2000). *The emergence of the speech capacity*. Mahwah, NJ: Lawrence Erlbaum Associates.

Oller D. K., & Eilers, R. E. (1988). The role of audition in infant babbling. *Child development*, *59*(2), 441-9.

Oller, D.K., & Lynch, M.P. (1992). Infant vocalizations and innovations in infraphonology: Towards a broader theory of vocal development disorders. In C. Ferguson, L. Menn & C. Stoel-Gammon (Eds.), *Phonological development: Models, research and implications* (pp. 509-536). Parkton, MD: York Press, Inc.

Papoušek, M., Papoušek, H., & Bornstein, M.H. (1985). The naturalistic vocal environment of young infants: On the significance of homogeneity and variability in parental speech. In T. Field & N. Fox (Eds.), *Social perception in infants* (pp. 269-297). Norwood, NJ: Ablex.

Petersen S. E., Fox, P. T., Posner, M. I., Mintum, M. A., Raichle, M. E. (1989). Positron emission tomographic studies of the processing of single words. *J Cogn Neurosci, 1*, 153–70.

Piaget, J. (1954). *The construction of reality in the child*. New York: Basic Books.

Reznick, J. S., Fueser, J. J., Fingado, C., & Straight, N. (1998). Parents can predict infant intentional behaviors, *Infant Behavior and Development, 21*, 64 (Supplement 1, Special ICIS Issue).

Riksen-Walraven, J. M. A. (1978). Effects of caregiver behavior on habituation rate and self-efficacy in infants. *International Journal of Behavioral Development, 1, 105-130.*

Smith, L. B. (2005) Action Alters Shape Categories. *Cognitive Science, 29*, 665-679.

Soska, K. C., Adolph, K. E., & Johnson, S. P. (2010). Systems in development: Motor skill acquisition facilitates 3D object completion. *Developmental Psychology*, *46*, 129-138.

Tamis-LeMonda, C. S., Adolph, K. E., Lobo, S. A., Karasik, L. B., Dimitroupoulou, K. D., & Ishak, S. (2008). When infants take mothers' advice: 18-month-olds integrate perceptual and social information for guiding motor action. *Developmental Psychology, 44*, 734-746.

Tao, S., & Dong, Q. (1997). *Referential gestural communication and locomotor experience in urban Chinese infants*. Unpublished manuscript, Beijing Normal University, Beijing, China.

Thelen, E. (1996). Normal infant stereotypies: A dynamic systems approach. In R. L. Sprague & K. M. Newell (Eds.), *Stereotyped Movements: Brain and Behavior Relationships*. (pp. 139-165). Washington, DC: American Psychological Association.

Thelen, E., & Smith, L. (1994). *A Dynamic Systems Approach to the Development of Cognition and Action*. Cambridge, MA: MIT Press.

Thiessen, E. D., Hill, E., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy, 7*, 53-71.

Werker, J. F., Pons, F., Dietrich, C., Kajikawa, S., Fais, L., & Amano, S. (2007). Infant-directed speech supports phonetic category learning in English and Japanese. *Cognition*, *103*(1), 147-162.

West, M. J., & Rheingold, H. L. (1978). Infant stimulation of maternal instruction. *Infant Behavior and Development, 1*, 205–215.

Yoshida, H. & Smith, L. B. (2005). Linguistic cues enhance the learning of perceptual cues. *Psychological Science*, *16*(2), 90-95.

CHAPTER 5

BEGGING FOR SONG: SOCIALLY MOTIVATED VOCAL LEARNING IN ZEBRA FINCHES

Abstract

We examined social motivation in song learning in the zebra finch, *Taeniopygia guttata*. Young males most commonly learn the song of their fathers, even when other potential tutors are available, unless the other tutors interact with and feed them (ref). In addition, young zebra finches are nutritionally dependent on their fathers during the beginning of the sensitive phase for song learning. Specifically, the father provides young with three things – food, interaction and song. Could both food and social interaction serve as motivating factors that become associated with the father and facilitate vocal learning? We made fine-grained behavioral observations of father-infant interactions in the period immediately after fledging and before the end of the sensitive phase of song learning. We focused on specific vocal behaviors that the father engaged in while feeding the young. We found that feeding events tend to co-occur with adult male song and the response of adult males to infant begging transitions from a higher proportion of feeding-responses to a higher proportion of song-responses from Day 25 to Day 30. Further, over this time the increase in proportion of paternal responses that were comprised of male song predicted song outcome measures at the time of song crystallization. We propose that early interactions with the father that center around feeding help juveniles associate the reward value of food with fathers' song. Thus young male zebra finches are motivated to learn the song of their fathers because the father is the one that feeds and maintains the greatest

amount of social interactivity with them. As in humans, a socially motivated system of learning guides vocal learning in zebra finches.

Introduction

Birdsong and human speech have long been considered parallel systems in the study of vocal learning [Doupe and Kuhl, 1999; Goldstein et al., 2003]. However, compared to investigations of the acquisition of acoustic structure, the role of social motivation in vocal learning remains understudied. Locke speculated that affective exchanges with the caregiver act as a passage-way for early language learning in human infants (Locke, 1993). Recent studies in humans have indicated that vocal learning is shaped by social motivation. Infants learn preferentially from conspecifics that engage in socially contingent interactions with them (Kuhl et al., 2003) and there is good evidence that reinforcement embedded in social contingency guides prelinguistic vocal learning (Goldstein et al., 2003, 2008). Other motivating features of social interactions in humans, such as the prosody of infant-directed speech, are also known to arouse infant attention and modulate infant affect (Fernald, 1985; other refs). Recent work in comparative developmental neuroanatomy has pointed at homologies in the embryological origins of human amygdala/basal forebrain limbic regions and avian song centers. These data suggest that bird song nuclei are more soundly positioned in regions that typify social and affective learning and regulation in mammals and underscore the role of social motivation in vocal learning in both humans and avian species (Syal & Finlay, 2010).

The idea that social motivation derived from social interactions organizes vocal behavior has received little attention in studies of song learning. Songbirds are typified by their ability and need to learn their species typical song from the environment. While some species, such as the

song sparrow(s), are able to learn from tapes, many require the presence of a socially interactive tutor and it has been argued that learning from tapes constitutes a type of social learning as well (Janik and Slater, 2000). Cross fostering experiments have shown that adult reinforcement of a subspecies dialect can motivate the learning of that dialect, different from the bird's own (Freeberg, King, & West, 2001). Indeed, heterospecific tutors can reinforce the learning of heterospecific songs altogether, even in songbirds that reject the audio-taped songs of alien species, (Baptista and Petrinovich, 1986)

Zebra finches are monogamous pair-bonding birds that display biparental care of young and live in large colonies (often ranging over a 100 birds) with individual families ranging from 4-6 members (Zann, 1996). Despite this large proliferation of potential song tutors, juvenile male finches tend to learn to produce the song of their fathers[4] (Mann and Slater, 1995). Similarly, juvenile female finches, who do not sing, display a preference for the father's song over the songs of other males (Riebel, 2000; Riebel et al. 2002).

Although the majority of studies have concluded that young male zebra finches learn their father's songs, there is some evidence to indicate that they may learn from other males as well in an aviary setting, leading to hybrid songs. In these cases, it was also noted infant finches learn from males that often engage in interactive resource allocation - such as feeding - with them [Williams, 1990]. Early cross-fostering experiments showed that young males preferentially learn the song of a heretospecific foster father (in this case, a Bengalese Finch), even when adult conspecific males could be heard nearby [Immelman. 1969]. This led Immelman to propose that infant finches prefer to learn the song of the male with whom they have a personal bond based on a provisioning relationship (Immelmann, 1969; Zann, 1997).

[4] Zebra finches learn a single stereotyped song.

While the sensitive period for song learning lasts up to day 65-80 in zebra finch males, young that are isolated from their fathers after day 35 are still able to reproduce his song in a manner akin to non-isolates (Bohner, 1990, Mann et al., 1991, Slater and Jones, 1995). Thus zebra finches likely learn the main tenets of their song between fledging and day 35 (but see Eales, 1985, 1987; discussion section of this paper). It has been noted that during this period they remain dependent on their parents for both resource allocation (food) and adult social influences (Zann, 1996). Thus the father normally provides young males with three things – food, interaction and song.

Could both food and social interaction serve as motivating factors that become associated with the father and thereby facilitate vocal learning? Here we studied the relationship between resource allocation and the availability of song models in the zebra finch. We made fine-grained measurements of social interactions between father and juvenile finches over developmental time All birds were housed with their families in order to approximate all of the social degrees of freedom they would typically have as closely as possible. Begging in zebra finches has previously been studied in the context of the effect of begging cost on fitness (Parker et al., 2002) and the impact of biparental versus uniparental care on begging and growth rate (Royle, Hartley and Parker, 2006). To the best of our knowledge, the relationship between parental responses to begging and song learning has never been examined prior to this study.

Our hypothesis was that young male zebra finches learn the song of their fathers because the father is the one that feeds them and maintains the greatest amount of social interactivity with them. We reasoned that specific features of the interactions of juvenile males with their fathers must endow his song with motivational significance. Specifically, we wanted to assay for temporal contingencies between feeding and the production of adult male song by the father.

Materials and Method

We made high-resolution behavioral observations of father-infant interactions in the period immediately after fledging and before the end of the sensitive phase of song learning. We focused on specific vocal behaviors that the father engaged in as he fed the young.

Subjects and Housing

The subjects used in this study are part of a larger (ongoing) study cohort of 13 families. Of the 13 families, recordings from the ages of interest i.e. Day 25 and Day 30 post hatching were only available for 10 families. Of the 10, the recording from 1 family did not contain any begging on day 30, and the recordings from another 2 families did not contain any begging on day 25. Therefore, 7 families of zebra finches were used for this study. The data from all the juveniles in any given family were pooled to obtain a family score for each of the variables. 5 families were used for the initiation and longitudinal analyses because the male juveniles from two of the families died prior to reaching crystallized song. Families, consisting of a parent dyad with 3-5 (M = 4) offspring, were housed in single-nest hatching (dimensions: 46 X 44 X 36 cm) that were stacked on top of each other, 3 per column. At the mean age of 23 days, the families were moved to a larger recording cage (dimensions: 62 X 43 X 36 cm), which were also similarly stacked. The cages were kept in community zebra finch housing which provided for visual and acoustic contact with birds from other families. Nest boxes were removed from the cages between days 33 and 36 (M = 35.29) The age range for the offspring used in this study is 25-26 days for Day 25 analyses and 30-31 days for the Day 30 analyses of the begging/response coding. Juveniles were kept with the family till an average age of 94.29 days (range: 84 – 99). The cages were kept in laboratory rooms maintained at 21°C on a 12:12 h light: dark cycle. Food and water in both conditions were available ad libitum.

Recording

Families placed overnight in a recording room which contained a sound-attenuating tent (size 140 X 98 X 80 cm) made of PVC pipe and sound attenuating foam. All recordings were done next morning between 8 am and 12 pm to control for activity of the birds. Each family was video recorded for 60 minutes per day after the offspring were fledged and recordings were spaced an average of 3.58 days apart. The data presented in this study were obtained from recordings made on day 25 and day 30. The recordings were made using a Sennheiser shot gun microphone attached to a Canon MiniDV ZR930 camcorder onto Fujifilm DV Cassette miniDVs at an audio resolution of 44.1 KHz. The tapes were ingested with a JVC Super VHS ET professional deck at 32 KHz rate for digitization and behavioral annotation.

Behavioral Coding

Familial interactions surrounding feeding and song were annotated using ELAN: EUDICO Linguistic Annotator Version 3.7.2 (Copyright © 2001 - 2009 Max-Planck-Institute for Psycholinguistics, Nijmegen, The Netherlands) which is also capable of frame by frame coding. Specific behaviors that were coded included begging, feeding and adult male song. Feeding or adults male song from the father was coded as a response to begging if it occurred within a 15 second contingency window following the termination of the begging bout. We also noted the identity (based on band color) of the juvenile who initiated each begging bout.

Begging: We scored begging behavior according to (Muller and Smith, 1978) wherein typical zebra finch begging involves the begging call, open mouth, tongue waggling and head/neck rolling and body squirming. Begging behaviors that were separated by a gap of less than 1 second were categorized as the same begging bout.

Feeding: A parent was scored as having fed a chick when it inserted its bill into the

chick's gaping mouth and could be seen regurgitating, with characteristic heaves of the body. Feeding duration lasted from the point at which the parent inserted its bill until it was withdrawn from the chick's mouth (Royle, Hartley and Parker, 2006)

Male song: Adult male zebra finch song consists of 1 individual specific phrase consisting of 3-14 elements learnt from the father and repeated, on average, 1 to 8 times (Slater et al. 1988; Zann, 1996). All instances of adult male/father song in each 60 minute were annotated. Song phrases separated by less than 1 second were annotated as part of the same song.

Song Quality

Mean juvenile age for recording of crystallized song was 138.7 days (range: 106-212). Directed song was recorded in a sound proof room. After the last day of family recordings (Mean = 95.5 days), the juveniles were put into same-sex aviaries with other inexperienced birds. They were placed in the sound-proof room the night before the day of crystallized song recording and on the day of recording, a female was placed in a neighboring cage in the room with them. Directed songs were recorded for 60 minutes, which usually yielded at least 20 clear (without background noise and female chirping) motifs. If the juveniles did not sing, the same procedure was repeated the next day.

Measures of song quality were obtained using Sound Analysis Pro (SA+) software. The SA+ software reduces song to set of five simple, one-dimensional measures: pitch, amplitude modulation, frequency modulation, Weiner entropy and goodness of pitch. Since each of these features has different units and statistical distributions, to arrive at an overall score of similarity, the units for each feature are transformed to units of statistical distances. Euclidean distances across all features can then be computed for frames (short units of time ~ 10 ms) and intervals (larger units of time ~ 70ms) (Tchernichovski and Mitra, 2000). Under the default settings,

"similarity sections" are neighborhoods of intervals in the similarity matrix computed by SA+ wherein the p value of Euclidean distance is less than 5% for all neighbors. The actual similarity values are calculated frame-to-frame within the similarity section. SA+ computes similarity at local (frame level) as well as at the section level. To obtain a unique similarity estimate, SA+ eliminates redundancy by cutting out sections that overlap with other sections that explain more similarity. As the last step, SA+ computes *final sections*, which are superior similarity sections that pass a final stage filtering in the analysis process by omitting similarity sections that explain very little similarity (and are likely to be noise).

The overall combined similarity score computed by SA+ is a product of 3 components: % similarity, accuracy and sequential match.

Overall combined score = Percentage similarity * accuracy * sequential match.

Percentage similarity: is the percentage of tutor's sounds included in *final sections.*

Accuracy: The average local similarity score across *final sections*, accuracy is quantification of the vocal match of acoustic features between the two songs, assuming that the sounds are related.

Sequential match: Sequential match is calculated by sorting the *final sections* (as laid out in the SA+ manual*)* according to their temporal order in reference to sound 1, and then examining their corresponding order in sound 2.

For the purposed of our study, three individual scores as laid out above: accuracy, percentage similarity and sequential match, as well as the overall combined similarity scores was calculated for each juvenile male in an asymmetric comparison wherein song 1 was tutor (father) song and song 2 was juvenile song at crystallization. The first 10 clear motifs from the fathers' song recordings were compared to the first 20 clear motifs from juveniles. For our analyses, we

used the default settings of the Sound Analysis Pro similarity module.

Results

In order to study the transition between begging and learning to sing, we decided to focus on father-infant interactions at two time points during the sensitive period for song learning, day 25 and day 30. This decision was motivated by the desire to examine the later period of begging because it is proximal to the onset of sub-song in developmental time.

Is the amount of begging different between days 25 and 30?

As a first step, we tabulated the total number of begging bouts and the changes in begging over time. A paired samples T test indicated that the amount of begging across families is not significantly different between day 25 (M = 12.00, SD = 6.08) and day 30 (M = 12.14, SD = 7.53), $t(6) = -0.079$, $p = .93$ (Figure 5.1).

Who do the juveniles beg from?

We then assessed the different proportion of begging directed at the father and the mother. Begging was aggregated over all the juveniles in each family. A paired T test revealed that the overwhelming majority of begging is directed at the father (M = 0.85, SD = 0.19) with one a very small percentage of total begging at this age being directed at the mother (M = 0.15, SD = 0.189), $t(6) = 4.89$, $p = .003$ (Figure 5.2).

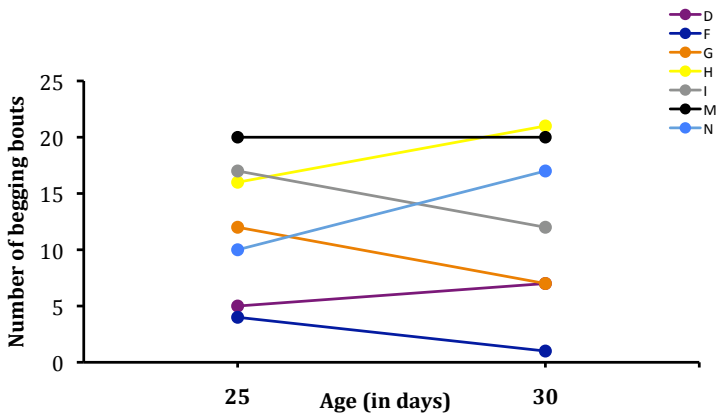

Figure 5.1: Changes in total amount of begging between the two ages: The amount of begging is not significantly different across families, p = 0.93

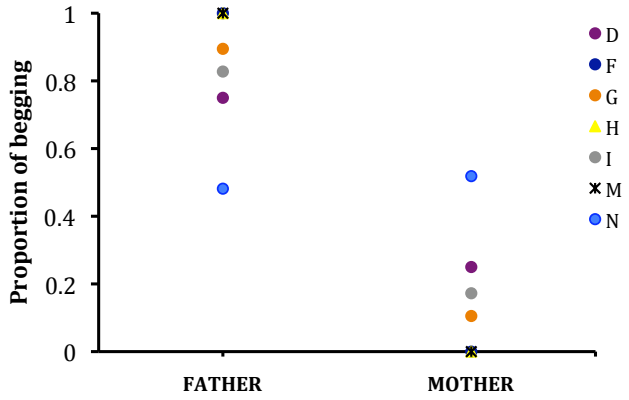

Figure 5.2: Proportion of total begging directed at individual parents. Juveniles beg significantly more from the father, *p = 0.003

What does the father do in response to begging?

We annotated feeding and song behavior during begging bouts and in the 15 seconds immediately following each bout. A larger proportion of begging bouts were responded to by feeding at Day 25 (Mdn = 0.5) than at Day 30 (Mdn = 0.05), T = 0, p < .025, r = .88), by which time, a change occurred in the fathers' behavior. A Wilcoxon Signed-ranks test indicated that a significantly larger proportion of begging was responded to by song at day 30 (Mdn = 0.5) as compared to day 25 (Mdn = 0.25), T = 1, p < .025, r = .82. Both feeding and singing within the

15 sec period following each begging bout also constituted a small proportion of the fathers' response. While we were unable to conduct a significance test because not all fathers showed this response at both ages, we noted that feeding always preceded singing when the father responded with both food and song.

Can the link between begging and male song be explained by unrelated changes in singing?

This change in the father's response could also be explained by chance pairing between begging and song and/or increased singing from the father as the juveniles get older. To rule out the first possibility, we compared the proportion of singing during/after begging to the proportion of singing during randomly selected control intervals that were matched to the begging intervals for frequency and duration. Each 60 minute recording session was divided into 240 15-second bins. A random number generator was used to generate a number of onset times equivalent to the total number of begging bouts for any given family. The 15-second bin that started at the randomly determined onset time was then scored for the presence of adult male song. The proportion of begging that received a song response is significantly higher than the proportion of song that would be expected to occur by chance across the 7 families at both age = 25 days, (M_{obs} = .2886, SD = .215) (M_{rand} = .0843, SD = 0.093), t(6) = 3.81, p = .009 as well as age = 30 days (M_{obs} = .61, SD = .258) (M_{rand} = .11, SD = 0.094), t(6) = 4.01, p = .007. A directional Wilcoxon Signed-ranks test indicated that a significantly larger proportion of begging was responded to by song (Mdn =_0.52) than could be expected by random chance (Mdn = 0.1) in each of the families, T = 0, p < .01, r = .88 (Figure 5.3).

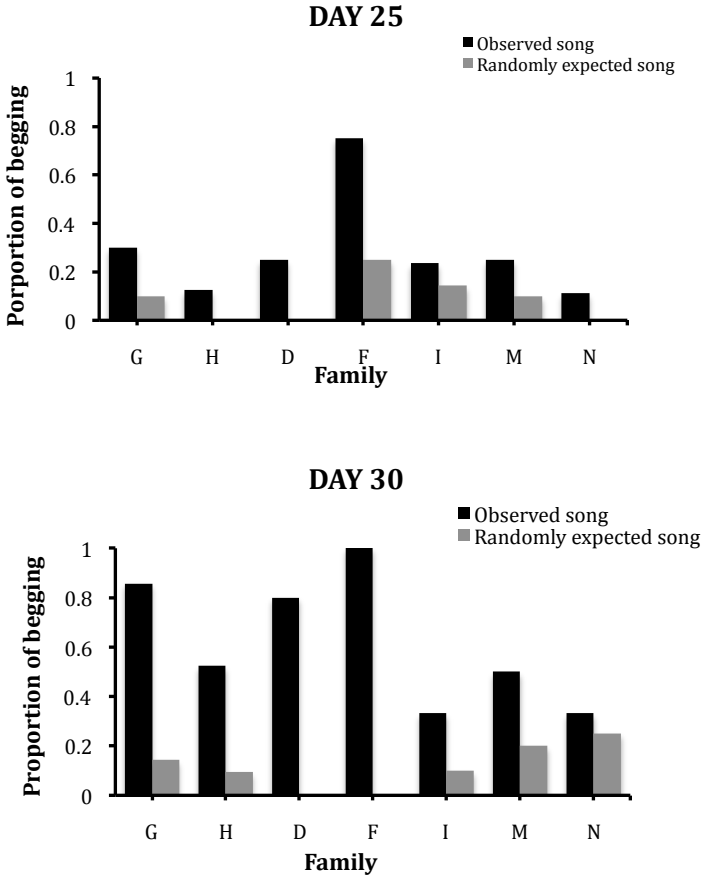

Figure 5.3: The observed proportion of begging that received a song response is greater than what would be expected by random chance on both days 25 and 30, *ps < .009.

Alternatively, the shift to a link between begging and male song at day 30 might be explained simply by increased singing by the adult male at that age. To address this question we tabulated the total number of male songs that occurred during our each of our recordings at day 25 (M = 19.85, SD = 17.87) and day 30 (M = 25.42, SD = 15.56). The overall number of songs per 60 minute sample did not differ across the two different time points $t(6)$ = - 1.041, p = .338 (Figure 5.4)

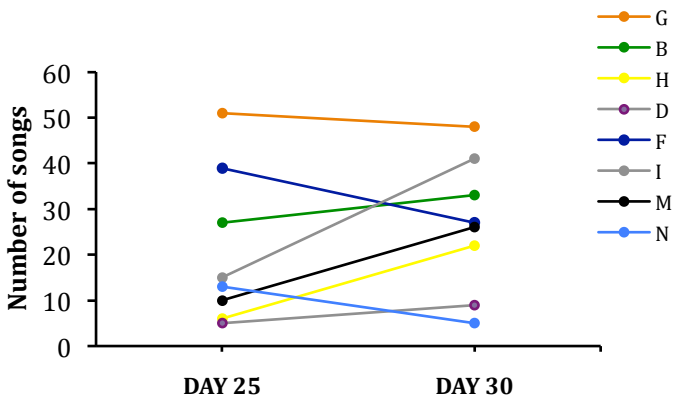

Figure 5.4: The overall number of songs that the fathers sang did not differ across the two ages, $p = .338$

Does paternal responsiveness have longitudinal effects on song quality?

A stepwise regression analysis indicated that the difference in proportion of begging (M = 0.293, SD = 0.10) that received male song between days 25 and 30 is significantly correlated with the combined score (M = 49.39, SD = 22.55) of song quality, $R^2 = .952$, $p < .001$ (Figure 5.5A). Family membership alone was not a significant predictor of song outcome $R^2 = .266$

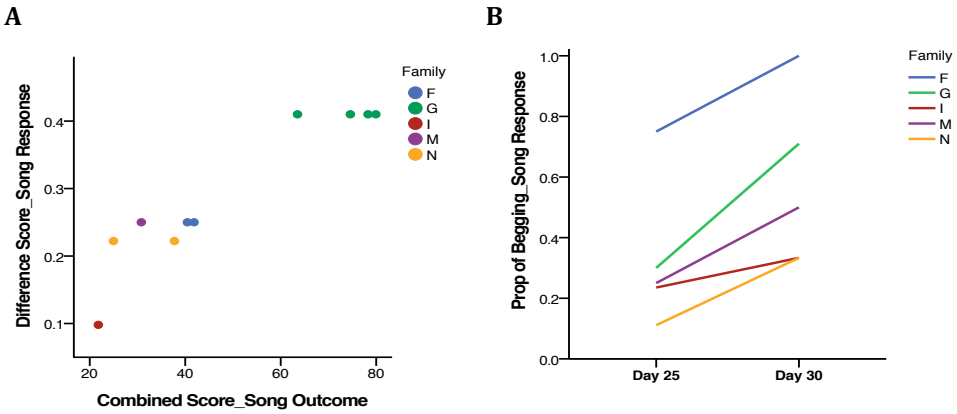

Figure 5.5: The difference in proportion of begging that received a song response between day 25 and day 30 is a predictor of song outcome at crystallization. A) Difference scores of all juvenile males from each family correlated with combined song outcome score, $R^2 = .952$*** B) Change in the proportion of paternal response that was comprised by song in each of the 5 families.

Discussion

After fledging, juveniles beg for food preferentially from the father. The father's responses to begging include feeding and song. The proportion of begging bouts that receive song rises dramatically as juveniles grow older and enter the sensitive period of song learning. The mean proportion of begging bouts across 7 families that received a song response was only 28% on day 25 as opposed to 51% on day 30. Co-occurrences between begging and song cannot be explained by a chance association between begging and song or by increases in paternal singing between the two ages. Additionally, paternal responsiveness as a whole does not differ between the two ages across families, indicting that the increases in adult male song in response to begging constitute a specific and directed change in the form of father infant interaction over early development. We propose that the fathers' responses to begging link the motivational value of feeding with a song model over the course of early vocal development in juvenile zebra finches. Further, fathers who make a bigger change in the their proportion of song-responses to begging during the early part of the sensitive phase for song learning, rear sons who sing better songs as adults. It is noteworthy that it's the difference score between proportion of begging that receives a song-response at the two age points that matters and not the overall amount of responding. Thus fathers who make more significant changes in the form and patterns of their responses likely endow their song with greater motivational value, causing it to be better learnt by juveniles. While we only have the data for the juvenile males form this cohort, it would be interesting to see if the female-siblings of these males show a similar greater preference for the father's song.

These data suggest that the father's song holds motivational significance for juvenile finches due to its association with food reward. The value of food as a fundamentally rewarding

stimulus has been well documented across all studied species and acts through the dopaminergic reward pathway in the brain (Wise, 2006). Dopamine is believed to act as a "reward-alerting" signal that enhances the learning of appropriate behavioral reactions to the prediction of upcoming reward (Schultz, 2002). Recent evidence has shown that dopaminergic neurons within the VTA (ventral tegmental area) SNc (Substantia Nigra) and PAG (Peri-aqueductal gray), are all activated in juveniles (d35) in response to tutor song (Nordeen et al., 2009). All of these areas are involved in dopaminergic reward processing and are already known to respond to a number of song related social interactions (Hara et al., 2007, 2009). Additionally, Nordeen et al. (2009) found an enhanced effect of familiar tutor song (as opposed to unfamiliar song) on VTA and PAG, leading the authors to propose that the association of dopaminergic reward with striatally mediated synaptic plasticity facilitates the learning of tutor song in zebra finches. Earlier studies from the same group have shown that exposure to familiar tutor song at day 35 leads to phosphorylation of Ca 2+/calmodulin dependent protein kinase (CaMKII) in Area X of juvenile zebra finches. Autophosphorylation of CaMKII is critical for synaptic changes that underlie learning and developmental plasticity (Lisman et al., 2002; Colbran and Brown, 2004). Intriguingly, the same neurons within the striatum that express CaMKII also express DARPP-32, a dopamine signaling protein (Hein et al., 2007), again pointing to the idea that learning of a specific tutor's song is mediated through dopaminergic reward. Our results indicate that association with food motivation from the father might catalyze the early establishment of song as a rewarding stimulus.

The temporal contingency between juvenile begging and paternal song also fits in well with numerous studies that have used song for operant conditioning Operant conditioning in zebra finches is based on the reinforcing properties of song; conspecific song is reinforcing to

zebra finches (Braaten and Reynolds, 1999) and song reliably reinforces the behavior that precedes it (Stevenson-Hinde, 1973; Adret 1993 ; Houx and ten Cate 1999; Riebel, 2000; Ritschard, Riebel & Brumm, 2010). ten Cate (1991) found that adult males visit a perch that more frequently when it gives them exposure to song (from a tape) and also show a higher preference for the reinforced perch than yoked controls. However, the songs that these birds learn (in the absence of any other live tutor input) remain impoverished and resemble neither the taped song, nor the songs of their fathers. Thus the typical social environment that finches grow in facilitates song learning in additional ways. Our findings suggest that the reinforcing properties of song are established via a food reward-song linkage in early development that is shaped by father-juvenile social interactions.

In our experiment, the juveniles in each family consisted of both males and females. Female finches are known to show a preference for song that is similar to their fathers' (Clayton, 1988; Riebel et al., 2002). Our results provide a potential mechanism that could underlie the establishment of this preference. It is possible that juvenile female preference gets established in much the same way as juvenile males; females might prefer their father's song because it is associated with the motivational value of a food reward.

Early work with zebra finches showed that if males are removed from their fathers (and all other sources of tutorship) at day 35, they develop a highly abnormal song whereas if the separation is made after day 65, the father's song that was learnt is retained. If the separation is made at day 50 and a new potential tutor is provided, then a hybrid song is learnt (Eales, 1985). Eales had suggested that this pattern results from an early priming effect of the father's song (before day 35) but that critical elements of the song are learnt later in development. In another study, the same author found that males who were raised by their mothers and then introduced to

their fathers at day 35, 50 or 65 learnt varying amounts of their fathers songs with only the day 35 group learning 100 % of the fathers song and the other two groups learning 70% and 75% of their songs from the father (Eales, 1987b). An additional but separate body of data from the same subject cohort has indicated that counter-singing between fathers and juveniles wherein the father's song follows juvenile song reinforces juvenile song and/or correlates with song outcome measures. Further, wingstrokes and fluff-ups from the mother in response to juvenile song are also correlated with song outcomes measures at crystallization (Menyhart et al., 2009). Therefore, it appears that no one aspect of parental behavior is responsible for song learning. Instead, vocal learning in the zebra finch is coordinated through social motivation derived from the interactions with both parents, throughout the development to independence. Thus our data complement Eales's (1985, 1987) findings – we suggest that the priming effects of father's might be set up through food related interactions between fathers and juveniles early in development and further crystallization of specific song elements is achieved through continuing social interactions with conspecifics.

An idea that is central to the motivations driving this study is that the vocal learning that characterizes a species occurs in the context of its specific developmental niche. Learning in an ecological context arises from the interplay of environmental affordances and the learner's capacity to perceive such affordances [Gibson and Pick, 2000]. Socially motivated influences on vocal learning illustrate the manner in which the developmental niche is structured. Under normal circumstances, the adult males who are *likely to be present* within the sphere of learning influences for infant males, and are in addition *likely to acquire motivational valence* for the infant males via resource allocation and social interaction, are *likely to be their fathers*. Thus learning of the father's song is not indicative of an "innate" preference, but is instead indicative

of structure (the presence of fathers who will feed and interact with the young) that can be *expected* from the environment. We argue that the fact that in zebra finches, final termination of the plastic phase of song learning is not determined by the completion of the song learning process, but by the extent of social interactions with live tutors, supports this ecological framework of song learning. If birds are held in isolation in sound-proof chambers, then termination of the "sensitive" phase is delayed and these birds are capable of learning in adulthood when a tutor becomes available [Morrison and Nottebohm, 1992; Slater et al., 1993]. In accordance with the idea that rearing treatment can prolong the period of song learning from a tutor, rearing infant males in aviaries from which adult males have been removed, results in a preservation of a high number of dendritic spines on neurons of the LMAN (lateral magnocellular nucleus of the anterior nidopallium, a vocal song nucleus in the avian telencephalon) for 3 weeks longer than controls [Wallhausser-Franke et al. 1995]. In normal zebra finches, the number of dendritic spines per unit length in the LMAN decrease by over 50% between day 35 and 95 [Nixdorf-Bergweiler et al. 1995]. These data indicate not only that the presence of adult male tutors is necessary for normal song learning but also that there are some social experiences that are *expected* to be present normal early rearing environment. In the absence of these expected experiences, the developing system retains its plasticity – and this likely conferred an evolutionary advantage to these birds (West-Eberhard, 2003)

Our results indicate that social motivation drives early vocal learning in zebra finches. This is in agreement with studies of socially guided learning (Goldstein & Schwade, 2009) in songbirds as well as humans wherein positive reinforcement from conspecifics facilitates the development of more advanced forms. Female brown-headed cowbirds (*Molothrus ater*) provide silent feedback that shapes the development of the males' song. Silent contingent feedback from

caregivers can shape prelinguistic vocal development in human infants [West and King, 1988; Goldstein et al. 2003]. Female cowbirds selectively respond to more advanced vocalization and thereby facilitate the transition to a developmentally advanced direction [West and King, 1988]. Similarly, human caregivers are more likely to respond to more advanced prelinguistic vocalizations that contain more speech-like elements [Gros-Louis, West, Goldstein, & King, 2006]. In cowbirds, males that receive a wingstroke are significantly more likely to repeat the song elements they had sung right before the wingstroke [West and King, 1988].

Since we were only able to make recordings of familial behavior post fledging, it is currently unclear how early in development this link is first made. All of the fathers we studied responded to some proportion of begging with song even at day 25, highlighting the possibility that the change in response from food to song might be even more gradual, making the father's song even more salient. Indeed, this does seem to be the case for the 1 family from our study cohort for which we have data at 5 different ages between Day 19 and Day 30 (Figure ref). However, additional data is obviously needed before any real claims regarding the nature of this developmental progression can be made. Further, we don't know if the pattern of primarily paternal responding to begging holds while the chicks are still nestlings. Zann (1996) had suggested that mothers are responsible for a larger part of parental care while chicks remain in the nest. It would be interesting to note if maternal provisioning in the nest bears any similar temporal contingency with male song from the father as or after the mother feeds the young even earlier in development.

We propose that in zebra finches, early interactions with the father that center around feeding help juveniles to make the association between the reward value of food and song that is delivered via the same source i.e. the father. Thus early interactions between the father and

juveniles create opportunities for the song to gain motivational power. These data provide evidence that socially situated motivation drives the learning of species typical vocal forms in zebra finches.

REFERENCES

Adret, P. (1993). Operant conditioning, song learning and imprinting to taped song in the zebra finch. *Animal Behavior. 46*, 149–59.

Airey, D. C. & DeVoogd, T. J. (2000) Greater song complexity is associated with augmented song system anatomy in zebra finches. *Neuroreport* 11, 2339–2344.

Buchanan, K. L., Spencer, K. A., Goldsmith, A. R., and Catchpole, C. K. (2003) Song as an honest signal of past developmental stress in the European starling (*Sturnus vulgaris*) *Proc. R. Soc. Lond. B* (270), 1149-1156

Buchanan, K. L., Leitner, S., Spencer, K. A., Goldsmith, A. R., and Catchpole, C. K. (2004) Developmental stress selectively affects the song control nucleus HVC in the zebra finch. *Proc. R. Soc. Lond. B* (271), 2381-2386

Baptista, L. F., & Petrinovich, L. (1986). Song development in the white-crowned sparrow: Social factors and sex differences. *Animal Behavior*, *34*, 1359–1371.

Bohner, J. (1990). Early acquisition of song in the Zebra Finch. *Taeniopygia guttata. Animal Behavior*, 39, 369-74.

Braaten, R.F., & Reynolds, K. (1999). Auditory preference for conspecific song in isolation-reared zebra finches. *Animal Behaviour*, *58*(1), 105-111.

11.Youngmaleswereraisedbytheirmothers(noadult male present) until they were 30 days old.

O. Tchernichovski, P. P. Mitra, Sound Analysis User Manual, version 1.01 (ed. 2, 2000) (available at http://talkbank.org/animal/sa.pdf).

Clayton, N. S. (1988). Song discrimination learning in zebra finches. *Animal Behaviour, 36*, 1016-1024.

Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: common themes and mechanisms. *Ann. Rev. Neurosci., 22*, 567-631.

Eales, L. A. (1985). Song learning in zebra finches: some effects of song model availability on what is learnt and when. *Animal Behaviour*, *33*(4), 1293-1300.

Eales, L. A. (1987a). Do zebra finch males that have been raised by another species still tend to select a conspecific song tutor? *Animal Behaviour, 35*(5), 1347-1355.

Eales, L. A. (1987b). Song learning in female-raised zebra finches: another look at the sensitive phase, *Animal Behaviour*, *35*(5), 1356-1365.

Eales, L. A. (1989). The influences of visual and vocal interaction on song learning in Zebra finches, *Animal Behaviour, 37*(3), 507-508.

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behav. Dev., 8*(2), 181-195.

Freeberg, T. M., King, A. P., & West, M. J. (2001). Cultural transmission of vocal traditions in cowbirds (Molothrus ater) influences courtship patterns and mate preferences. *Journal of Comparative Psychology*, *115*, 201–211.

Gibson, E. J., & Pick, A. D. (2000). *An ecological approach to perceptual learning and development*. Oxford: Oxford University Press.

Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proc. Nat. Acad. Sci., 100*(13), 8030-8035.

Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychol. Sci., 19*(5), 515-523.

Goldstein, M. H., & Schwade, J. A. (2009a). From birds to words: Perception of structure in social interactions guides vocal development and language learning. In M. S. Blumberg, J. H. Freeman, & S. R. Robinson (Eds.), *The Oxford Handbook of Developmental and Comparative Neuroscience*. Oxford University Press.

Gros-Louis, J., West, M. J., Goldstein, M. H., & King, A. P. (2006). Mothers provide differential feedback to infants' prelinguistic sounds. *International Journal of Behavioral Development*, 30(6), 509-516.

Holveck, M. J. & Riebel, K. (2007). Preferred songs predict preferred males: consistency and repeatability of zebra finch females across three test contexts. *Animal Behaviour*, *74*(2), 297-309.

Hara, E., Kubikova, L., Hessler, N. A., & Jarvis, E. D. (2007). Role of the midbrain dopaminergic system in modulation of vocal brain activation by social context. *Eur. J. Neurosci*., *25*, 3406–3416.

Hara, E., Kubikova, L., Hessler, N. A. & Jarvis, E. D. (2009). Assessing visual requirements for social context-dependent activation of the songbird song system. *Proc. Biol. Sci., 276*, 279–289.

Hein, A. M., Sridharan, A., Nordeen, K. W. & Nordeen, E. J. (2007). Characterization of CaMKII-expressing neurons within a striatal region implicated in avian vocal learning. *Brain Res*., *1155*, 125–133.

Houx, B.B., & ten Cate, C. (1999). Learning from playback tutor song in zebra finches: is there an effect of operant contingency? *Animal Behaviour 57*: 837–845.

Immelmann, K. (1969). Song development in the zebra finch and other Estrildid finches. In R. A. Hinde (Ed.), *Bird Vocalizations* (pp. 61-74). Cambridge: Cambridge University Press.

Janik, V. M., & Slater, P. J. B. (2000). The different roles of social learning in vocal communication. *Animal Behaviour*, *60*(1), 1-11.

Levrero, F., Durand, L., Vignal, C., Blanc, A., & Mathevon, N. (2009). Begging calls support offspring individual identity and recognition by zebra finch parents, *Comptes Rendus Biologies, 332*(6), 579-589.

J. Lisman, J., Schulman, H., Cline, H. (2002). The molecular basis of CaMKII function in synaptic and behavioural memory, *Nature Reviews Neuroscience*, 3, 175-190.

Locke, J. (1993). *The Path to Spoken Language*. Cambridge, MA: Harvard University Press.

Mann, N. I., Slater, P. J. B., Eales, L. A., & Richards, C. (1991) The influence of visual stimuli on song tutor choice in the zebra finch, Taeniopygia guttata, *Animal Behaviour*, *42*(2), 285-293.

Mann, N. I., & Slater, P. J. B. (1995). Song tutor choice by zebra finches in aviaries, *Animal Behaviour*, *49*(3), 811-820.

Morrison, R.G. & Nottebohm, F. (1992). Role of a telencephalic nucleus in the delayed song

learning of socially isolated zebra finches. *Journal of Neurobiology*, *24*, 1045-64.

Muller, R. E. & Smith, D. G. (1978). Parent–offspring interactions in zebra finches, *Auk 95*, 485–495.

Nixdorf-Bergwieler, B. E., Wallhausser-Franke, E., & DeVoogd, T. J. (1995). Regressive development in neuronal structure during song learning in birds. *J Nuerobiol*, *27*, 204-215.

Nordeen, E. J., Holtzman, D. A., & Nordeen, K. W. (2009). Increased Fos expression among midbrain dopaminergic cell groups during birdsong tutoring. *European Journal of Neuroscience, 30*(4), 662-670.

Nowicki, S., Peters, S. & Podos, J. 1998 Song learning, early nutrition and sexual selection in songbirds. *Am. Zool*. 38, 179–190.

Menyhart, O., Goldstein, M. H., & DeVoogd, T. J. (2009). Social Shaping in the zebra finch: Timing of interactions with adults affect song learning. Presented at the 42nd *Annual Meeting of the International Society for Developmental Psychobiology,* Oct 14-17, 2009, Chicago.

Parker, G. A., Royle, N. J. & Hartley, I. R. 2002. Intrafamilial conflict and parental investment: a synthesis. *Philosophical Transanctions of the Royal Society of London, Series B*, 357, 295-307.

Riebel, K. (2000). Early exposure leads to repeatable preferences for male song in female zebra finches. *Proc. R. Soc. Lond. B*. 267, 2553-2558.

Riebel, K., Smallegange, I. M., Terpstra, N. J., & Bolhuis, J. J. (2002). Sexual equality in zebra finch song preference: evidence for a dissociation between song recognition and production learning. *Proc R Soc B*, *269*, 729-733.

Ritschard, M., Riebel, K., & Brumm, H. (2010). Female zebra finches prefer high-amplitude song, *Animal Behaviour*, *79*(4), 877-883.

Roper, A. & Zann, R. (2006). The onset of song learning and song tutor selection in fledgling zebra finches. *Ethology, 112*, 458-470.

Royle, N. J., Hartley, I. R., Parker, G. A. (2006). Consequences of biparental care for begging and growth in zebra finches, Taeniopygia guttata. *Animal Behaviour*, *72*(1), 123-130.

Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron*, *10, 36*(2), 241-63.

Slater, P. B. J., Jones, A. E., & ten Cate, C. J. (1993). Can lack of experience delay the end of the sensitive phase for song learning? *Netherlands Journal of Zoology*, *40*, 80-90.

Slater, P. J., Eales, L. A., & Clayton, N. S. (1988). Song learning in zebra finches (*Taeniopygia guttata*): progress and prospects. *Adv. Study Behav., 18*, 1–33.

Slater, P. J. B., & Jones, A. E. (1995). The timing of song and distance call learning in zebra finches. *Animal Behaviour*, *49*(2), 548-550.

Stevenson-Hinde, J. G. (1973). Constraints on reinforcement. In R. A. Hinde & J. G. Stevenson-Hinde (Eds.), *Constraints on Learning* (pp. 285-296), London: Academic Press.

Syal, S. & Finlay, B. L. (*in press*). Thinking outside the cortex: Social motivation in the evolution and development of language, *Developmental Science*.

ten Cate, C. (1991). Behaviour-contingent exposure to taped song and zebra finch song learning. *Animal Behaviour*, *42*(5), 857-859.

Wallhausser-Franke, E., Nixdorf-Bergwieler, B. E., & DeVoogd, T. J., (1995). Song isolation is associated with maintaining high spine frequencies on zebra finch LMAN neurons. *Neurobiol. Learn. Mem*, *64*, 25-35.

West-Eberhard MJ. *Developmental plasticity and evolution.* Oxford: Oxford University Press; 2003.

West, M. J., & King, A. P. (1988). Female visual displays affect the development of male song in the cowbird. *Nature*, *334*(6179), 244-246.

Williams, H. (1990). Models for song learning in the zebra finch: fathers or others? *Animal Behaviour*, *39*(4), 745-757.

Wise, R. A. (2006). Role of brain dopamine in food reward and reinforcement. *Philos. Trans. R. Soc. Lond. B Biol. Sci., 361*, 1149–1158.

Zann, R. (1997). Vocal learning in wild and domesticated zebra finches: Signature cues for kin recognition or epiphenomena? In C.T. Snowdon, & M. Hausberger (Eds.) *Social influences on Vocal Development*. Cambridge, UK: Cambridge University Press, 86-97.

Zann, R. E. (1996). *The Zebra Finch: Synthesis of Field and Laboratory Studies*. New York: Oxford University Press.

CHAPTER 6

CONCLUSIONS AND DISCUSSION

My thesis started with the assertion that language is a social construct and that this derives from the fact that the motivation to communicate is contingent on the presence of social conspecifics to communicate with. Further, it made the claim that social motivation is necessary for the development of language and that a link between sub-cortical social reward and cortical learning systems might have been the crucial evolutionary adaptation that led us to language. If this is true, it follows that socially situated motivation should facilitate species-typical vocal learning

Chapter 2 of this thesis laid out the theoretical framework that forms the basis for the empirical chapters that follow. Using a comparative approach, both neuroanatomical and behavioral data were presented from humans and birds. The ontogeny of avian song learning has long served as a model system for human language (Doupe and Kuhl, 1999; Goldstein et al., 2003). The proposed homology between most of the avian song nuclei and mammalian lateral pallium (amygdala/limbic regions), along with behavioral data from studies of song learning, provide evidence that motivation and vocal learning remain intrinsically linked in birds. Recent evidence from humans about the importance of social motivation in language learning reinforces this premise in human vocal learning as well.

Chapters 3 and 4 probed the role of social motivation in early vocal learning in human infants. Chapter 3 detailed a study that extrapolated the role of social motivation to phonological learning and found that socially contingent input from caregivers helped infants learn the sounds

of their language. Chapter 4 found that changing maternal motivations feeds back into the dynamic of mother-infant interaction and facilitated the learning of more advanced sounds.

Chapter 5 tested the role of social motivation in vocal learning in a songbird species, the zebra finch in a more targeted manner than is currently possible in humans and concluded that young zebra finches learned the songs of their father because they associated song with food reward from him.

Chapter 3: The role of social motivation in phonological learning:

Investigations of both human (Goldstein et al., 2003; Kuhl, Tsao, & Liu, 2003; Goldstein & Schwade, 2008a) and animal communication (Beecher & Burt, 2004; King, West, & Goldstein, 2005) have shown that contingent feedback from conspecifics constitutes a socially guided system of vocal learning. In the brown-headed cowbird, *Molothrus ater,* female cowbirds provide silent feedback that shapes the development of the males' song. Females selectively respond to more advanced vocalization and thereby facilitate the transition to a developmentally advanced song (West and King, 1988) Similarly, silent contingent feedback from caregivers shapes prelinguistic vocal development in human infants (Goldstein, King, & West, 2003). Additionally, human caregivers are more likely to respond to more advanced prelinguistic vocalizations that contain more speech-like elements (Gros-Louis, West, Goldstein, & King, 2006) Thus the data from both of these vocal learning systems indicate that the development of the adult repertoire of song(s) or speech is a socially distributed system of learning wherein social reinforcement from conspecifics facilitates the development of more advanced forms.

Specific forms of contingently delivered vocal feedback cause infants to change correspondingly specific features of their own vocalizations towards developmental

advancement (Goldstein and Schwade, 2008a). In this study, the form of mothers' responses influenced infant vocal behavior. When mother responded with more resonant sounds, infants produced more fully resonant vocalizations themselves. When mothers responded with sounds that contained consonant-vowel (CV) articulations, infants increased the proportion of syllables with CV structure. This provided early evidence that infants are sensitive to the phonological patterns of the speech that they hear and are able to use social feedback driven learning to incorporate the phonology of ambient speech into their own vocalizations.

The canalizing effects of ambient adult speech have been well documented (Best, 1994). During the first year of life, infants develop perceptual categories that match the phonemic contrasts of their native language (Jusczyk, 1992; Werker &Tees 1984). Cross-cultural analyses have shown that both the vowels (Boysson-Bardies et. al, 1989) and the consonants (Boysson-Bardies & Vihman, 1991) that infants produce are similar to their ambient language. Pragmatic aspects of speech, such as the intonation patterns or the prosody that marks the babbling of French and English infants also resembles that of their ambient language by 11 months of age (Whalen, Levitt, and Wang, 1991).

Could socially motivated vocal learning be a mechanism that drives this change? Chapter 3 of this thesis addresses this question. If the answer to the above question is yes, then if infants from English speaking households are presented with exemplars from a novel phonological form in a socially motivated contingency paradigm, they should learn and incorporate these novel phonological contours in their own vocalizations. The babbling[5] of Nigerian infants commonly

[5] Babbling consists of a series of sounds that are developmentally immature and have not acquired all the infraphonological and phonological characteristics of the adult sound repertoire (Oller 2000). It has been hypothesized (Locke 1993) that babbling serves to maintain parental proximity and promotes social exchange. There is further evidence that feedback from the caregiver provides reinforcement of specific features of babbling that leads to the production of more mature and speech-like vocal forms (Oller, 1976; Goldstein et al., 2003). Infant babbling is a predictor of the

contains disyllabic vocalizations of the Vowel-Consonant-Vowel (VCV) form, as opposed to the Consonant-Vowel-Consonant-Vowel (CVCV) form typically seen in French infants (Boysson-Bardies, 1993). Thus utterances that follow the VCV phonology constitute a novel phonological contour for English speaking infants. In the experiments outlined in Chapter 3, we cued caregivers to respond to infant vocalizations with Yoruban VCV words while simultaneously leaning in, touching, and smiling at their infants. We also manipulated the variability of the input with one group of infants receiving a different Yoruban word for every maternal response and another group receiving the same word each time. We found that only infants in the contingent-variability group learnt and produced the novel phonological form. Further, their productions were non-imitative i.e. they extrapolated the underlying phonological rule – VCV – and incorporated it into their own novel VCV vocalizations.

While earlier studies have documented infants learning of statistical information from non-contingent presentations (Saffran et al., 1996; Marcus et al., 1999), it is important to note that these studies tested infant perception and not infant production of new phonological categories. Production of a novel acoustic form likely carries additional task demands over and above those of perception of the new categories.

An additional source of difficulty in our study might derive from language specific perceptual narrowing. 10-month old infants are not able to learn to discriminate non-native phonetic contrasts from the same distributional input that 6-8 month olds can (Maye et al., 2002, 2008); they need double the exposure/learning time to make perceptual sensitivity judgments about the same phonetic distinctions (Yoshida et al., 2010). Social feedback and reinforcement in our paradigm would aid learning by sourcing infant attention and making the new vocal forms

---

onset of first words, which is correlated with the number of different canonical syllable in an infant's repertoire. (Stoel-Gammon, 1992). Recent evidence indicates that babbling also serves as an arousal mechanism, enhancing sensitivity to perceptual information (Goldstein et al., 2010b).

more salient. Other work has shown that infants exposed to non-native speech contrasts show robust and durable learning of the contrasts only if they are presented by a interactive social agent but not if they are presented over video or audio (Kuhl et al., 2003) It was noted that infants in the social interaction group in the above study paid greater attention to their social counterparts than the audio-visual group infants (Conboy and Kuhl, 2007) and the predictive value of social factors (shared visual attention) in the learning efficacy of Spanish phonemes and words in English speaking infants has also been documented (Conboy et al., 2008) and thus *Chapter 3* provides evidence for the powerful effect that social motivation can exert on vocal learning.

There are multiple sources of social motivation embedded within the experimental paradigm used therein:

1. Contingent Vocal feedback:

Social stimulation contingent on infant vocalization has long been known to reinforce infant vocalizations rates, and to a greater extent than stimulation that is not contingent on infant vocal behavior (Poulson, 1983). Indeed, even a higher amount of non-contingent stimulation does not increase infant vocalization (Weisberg, 1963), highlighting the specific role of contingency.

In adult humans, the caudate nucleus within the dorsal striatum is sensitive to reinforcement of action and shows a robust response when subjects perceive a contingency between their actions and the outcome of the task (Tricomi et al., 2004). With specific regard to language learning, adult subjects who are given feedback during training show activation of the caudate (Tricomi et al., 2006) and the feedback brings about a dramatic boost in the learning of a non-native phonemic distinctions (McCandliss

et al., 2002). The caudate is implicated in affective and motivational processing shows a low incentive) with a significant correlation between feedback on high incentive trials and caudate activation (Delgado et al., 2004).

By the time infants are 5 month old, they have learnt the "instrumental value of babbling" and expect responses to their vocalizations (Goldstein et al., 2009b). Contingency and appropriateness of maternal responses are positively correlated with language outcome measures (Tamis-LeMonda, Bornstein and Baumwell, 2001; Rollins, 2003) and infants as young as 3 month old produce less nasalized (more mature) sounds in vocal turn taking interactions in which adults respond contingently to them (Bloom, Russel & Wassenberg; 1987; Masataka, 2003). It is easily conceivable that akin to adults, these effects of social contingency are accomplished through recruitment of striatal reward pathways in infants that act via modulation of infant motivation.

2. Physical reinforcement:

Within the social response period of the paradigm used in this experiment, mothers were instructed to "repeat the vocalizations they heard over their headphones *while getting closer to, touching and smiling at their infants*". Therefore, in addition to the reinforcing properties of vocal contingency, infants received 3 other socially motivating responses.

Approach and Touch: Research on the reinforcing properties of tactile stimulation goes back to Harlow (1958) who showed that given a choice between two surrogates, infant Rhesus Macaques who have been separated from their mothers show a preference for a cloth "mother" over a wire "mother". In humans, while vocal stimulation from adults has consistently proven to be the most potent reinforcer, only tactile stimulation is

also known to reinforce infant vocalization (Rheingold, Gerwitz and Ross, 1959; Staats & Staats, 1963; Haugan & Mcintire, 1972).

Pleasant touch is known to activate the orbitofrontal and cingulate cortices, as well as the striatum (Francis et al., 1999; Rolls et al., 2003). The OFC is believed to encode the reward value of appetitive stimuli (Rolls, 1999; Rolls, Burton and Mora, 1980) and shows activation is response to a wide range of both primary and secondary reinforcers such as positive and negative feedback (Elliott, Frith and Dolan, 1997), sex (Redoute et al. 2000), music (Blood et al., 1999) and money (O'Doherty et al., 2001).

In classroom education research, tactile stimulation has long been used in reinforcement therapy of children/adolescents with emotional/mental disorders (Bailey and Meyerson, 1969; Clements & Tracy, 1974a; Clements & Tracy, 1974b) Specifically, Clements & Tracy (1975) found that individual administration of tactile or verbal reinforcement resulted in higher levels of attention to task and task accuracy with a combination of both producing the highest effect. In an earlier experiment that followed the same social response paradigm as the current study, Goldstein et al (2003) asked caregivers to respond with only physical reinforcement i.e. approach, touch and smile without any vocal feedback. In the light of Clements & Tracy's (1975) findings, it would be interesting to compare the relative amount of vocal development between infants who received only physical reinforcement (Goldstein et al., 2003) and infants who received both vocal and physical feedback (Chapter 3; Goldstein and Schwade, 2008)

Tactile stimulation (stroking) has served as a reinforcer in classical conditioning tasks in both human infants (Sullivan et al., 1991) and rat pups (Sullivan & Hall, 1988). Given that approach and tactile stimulation remained closely associated with vocal

reinforcement within this experiment, hypothetically, infants could have learnt to view touch/tactile stimulation as a reward-predicting stimulus. The role of dopamine in behavioral reactions that are mediated by reward prediction as well as novel or salient stimuli (which the Yoruban exemplars could well be acting as) has been very well established (Schultz, 1999). It is also known that as the brain learns the association between reward and the reward-predicting stimulus, the dopaminergic response it transferred from the reward to the presentation of the stimulus (Schultz, 1998). The dopaminergic signal from the substantia nigra (SNc) and the ventral tegmental area (VTA) acts as global reinforcement for large parts of the striatum, the nucleus accumbens and the frontal cortex and can be has been referred to as a "reward-alerting signal" that activates numerous brain systems to learn "correct behavioral reactions to motivating environmental stimuli" (Schultz, 1999). Thus it is possible that in addition to their own reinforcing properties, maternal approach and touch (within our experimental paradigm) became motivating environmental stimuli that predicted reward in the form of vocal reinforcement. While it is clear that tactile stimulation was not sufficient to induce learning of the new vocal patterns in the absence of sufficient variability in the stimulus because of the absence of learning in infants in the contingent repetition group, it is nevertheless likely that given the necessary stimulus parameters in the environment (in this case variability), the motivational value of maternal touch facilitated the learning of the regularities of caregiver speech. A more detailed analysis of the temporal relationships between maternal touching and vocalizing would aid the assessment of the plausibility of this hypothesis.

Smiling: Smiling is a universal indicator of joy and is clearly a positively valenced affective signal in humans. By the time infants are 3 months old, they are able make distinctions between different facial expressions (Young-Browne et al., 1977; Field et al., 1982) and can discriminate smiling from other facial expression (La Barbera et al., 1976). They also show a preference for more intense smiles and it has been hypothesized that adult smiling may recruit infant attention (Kuchuk, 1986). 10 month olds are more likely to produce Duchenne (indicating reward) smiles (Duchenne, 1990 (1862)) when approached by their smiling mothers, highlighting the positive affect or motivational valence of the smile stimulus (Messinger, Fogel, & Dickson, 2001). In adults, smiling faces are rated as containing more positive affect than neutral faces (Messinger et al., 2001) and cause activation of the amygdala (Yang et al., 2002; Fitzgerald et al., 2006) perhaps the functional seat of motivation in the mammalian brain and involved in the perception of a variety of emotions, saliency, and biologically relevant information (Adolphs, 2008). Thus mothers' smiling as they touch and vocalized at this infant served as another source of social motivation within this study.

In sum, I would argue that the phonological learning displayed by the infants in this study occurred within a social reinforcement paradigm wherein multiple features of the social interaction between mothers and infants facilitated the learning and production of new distributional information. It has been suggested that human learning and memory is sensitive to the *recency* and the *frequency* of stimuli (Ellis, 2006) and cognitive processing is sensitive to both these independent variables additively: processing accuracy is described by a joint power law function of recency and accuracy (Schooler and Anderson 1997). The contingency of response in the experiments described in Chapter 3 coded for the *recency* of the stimulus,

possibly aiding language learning through the detection of co-incidences between input (maternal speech) and output (infant speech). The statistical variability of the input made the underlying phonological pattern most prominent and *frequent,* aiding the extraction and generalization of new vocal forms. And finally, social motivation salience-tagged the information embedded within the interaction, catalyzing learning.

Chapter 4: The role of locomotion in the social facilitation of vocal development

In the experiments outlined in chapter 4, we explored changes to infant vocal learning that are facilitated by changes in maternal motivation, as well infant (locomotor) ability. We found that infants' ability to locomote independently causes significant changes to maternal behavior, which in turn facilitate vocal learning.

The onset of independent locomotion is a seminal change in the infant developmental program and impacts multiple domains of early learning including perception, cognition and affect. It has been suggested that onset of self-produced locomotion is associated with increases in spatial cognition (Bertenthal et al., 1984, and Kermoian & Campos, 1988). Early evidence indicated that the development of depth perception, as evidenced by response to the "visual cliff" is related to crawling onset (Gibson and Walk, 1960) and the emergence of the emotion of fear with regard to the visual cliff is also related to locomotion (Bertenthal and Campos, 1990). Infants show greater sociability and affective displays after the onset of crawling (Campos et al., 1992) and there appears to be a reorganization of affective interactions between mother and infants upon the advent of independent walking (Biringen et al., 1995). The results of our experiment situate vocal learning within the domain of developmental processes that are affected by locomotion. In addition to increases in infant ability to interact with the environment, we

propose that the relationship between locomotion and vocal development is mediated by changes in maternal motivations and behavior.

In rhesus macaques, primiparous or experienced mothers encourage infants' independent locomotion in a manner that is sensitive to infant competence (Maestripieri, 1995) There is also some evidence that mothers distance themselves form infants and create opportunities for using the "pucker face" retrieval signal and its response in infants – which is to get proximal (Maestripieri, 1996). Simultaneously, there is also evidence that separation from infants in conjunction with removal to a novel environment recruits the HPA axis with significant cortisol release in mothers (Champoux & Suomi, 1993). While the exact nature of the effect of infant separation form the mother as a result of locomotion remains unclear, these experiments do indicate that primate mothers are cognizant of their infants' locomotor abilities and that the departure of/from the infant modulates their affective state. It seems likely that changes in motivational state in response to infant locomotion would bring about infant-directed behaviors from mothers.

In our experiment with humans, mothers increased the number of initiations they make towards the infants that involve both verbal and physical components when their infants were mobile and able to get distal from them. Many of these involved pointing at a toy or a picture within the playroom or handing it to the infant while naming it. A recent study in adults documented greatest amygdalar recruitment to novel (as opposed to familiar) stimuli and participants reported greatest arousal in response to the novel stimuli (Weierich et al., 2010). Amygdala activation is also associated with salience (Whalen et al., 2004) and memory enhancement (Kensinger & Schachter, 2006). Thus it is possible that mothers' initiations highlighting novel stimuli in the lab channeled the attention and motivation of their infants via

similar mechanisms of amygdala recruitment.

Mothers also increased the number of looks to infant in response to infant vocalizations. The role of eye gaze in vigilance-related arousal is well known and is used as an index for anxiety (Bradley at al., 2000; Mogg & Bradley, 2002). Thus increased looks directed at the infant when s/he was mobile, seem to provide credence to the idea that mothers of locomoting infants are more vigilant. Vigilance and attention remain intrinsically linked (Lim and Dinges, 2008) and increased attending to the infants causes increased maternal behavior directed at the infant, thus facilitating vocal development.

BOX 6.1 What's in it for the mum: Neural correlates of the motivational value of maternal behavior

Mothers provide infants with a highly responsive social environment that fosters learning. What motivates mothers to interact with their infants? The dopaminergic reward system of the brain seems to mediate maternal reward through its interactions with the prosocial neuropeptide, Oxytocin as well as opioid action within social/motivational circuits. Oxytocin causes an increase in prosocial approach behavior by dampening defenses (social avoidance) and activating reward in a manner that facilitates social attachment and affiliation (Young et al., 2001). One of the ways in which oxytocin exerts its effects on maternal behavior is through its action on the Ventral Tegmental Area (Pederson et al., 1994), a crucial part of the dopaminergic reward system (Alcaro, 2007) and also part of the social circuit). Injection of oxytocin antagonists into the VTA impair maternal behavior (Pederson et al., 1994). Direct infusion of oxytocin into the VTA increased the dopamine signal in the Nucleus Accumbens (Shahrokh et al., 2010) And there is increased dopamine release in the Nucleus Accumbens during mother-infant interaction (Hansen et al., 1993) – leading to the proposal that oxytocin mediates maternal reward from maternal behavior – which in turn is therefore re-inforced (Pederson, 1997).

Rat dams that show increased (licking/grooming) maternal behavior also have increased Oxytocin receptors in multiple nuclei of the social circuit* and administration of a oxytocin receptor agonist produces increased maternal behavior in rat dams that normally show significantly lesser licking/grooming of rat pups (Champagne et al., 2001)

In addition to vocalizing at infants, social interaction surrounding vocal learning includes looking, smiling, touching and listening to infants.

Looking and smiling: The rewarding affects of neoteny have long been appreciated and have more recently been demonstrated through direct activation of the nucleus accumbens, (within the mesocorticolimbic reward path) in response to

pictures of baby faces (Glocker et al., 2009) in nulliparous women. Another recent study found that the entire mesocorticolimbic dopamine reward pathway (midbrain VTA/SNc, striatum, and large parts of the frontal lobes including the prefrontal cortex, as well as the primary motor area) is activated in mothers in response to their own infant's face. Additionally, this reaction is heightened is the infant is smiling (Strathearn et al., 2008). The process of smiling is also believed to facilitates positive affect in the smiler (Soussignan, 2002) – so it is likely smiling at their infants modulates maternal affect and that this creates a motivated social environment for infant learners (as in Chapter 3). Amphetamines are among the most potent enhancers of vigilant attention (Magill et al, 2003; Cochran et al., 1992) and are believed to assert their arousing effects mainly through increases in dopaminergic transmission (Koob & Nestler, 1997; Leshner & Koob, 1999). It is possible that infant-stimuli-driven dopaminergic activation of vigilant attention mediated the changes in maternal behavior that we observed in Chapter 4.

Touching: The somatosensory stimulation of touch causes plasma oxytocin levels in rats to increase by 181% (Stock and Uvnas-Moberg, 1988). In humans, touching or contact with newborns causes Oxytocin release in mothers (Matthiesen et al., 2001) and plasma Oxytocin levels are correlated with touch, gaze, positive affect, and ID speech in the first month post partum (Feldman et al., 2007)

Listening: In rat pups, ultrasonic vocalizations from pups are very effective auditory retrieval cues in rats and mice (Smotherman et al., 1974) Similarly in humans, listening to (laughing and crying) infant vocalizations causes activation of brain regions associated with motivational processes such as the amygdala and the anterior cingulate cortex (Sander et al., 2007).

Chapter 5: Socially motivated vocal learning in Zebra Finches

Chapter 5 tested a direct prediction of the theoretical framework described in Chapter 2. On the basis of neuroanatomical and behavioral data from avians and humans, we hypothesized that vocal learning is modulated via affective feedback. Specifically, we argued that motivation derived from early interactions with social conspecifics drives vocal development.

Songbirds are defined by their ability to learn their song from the environment and constitute the great majority of species in the order Passeriformes. Birdsong is an elaborate vocalization that is distinct from other call vocalizations that are used by avian species in a variety of contexts such as parent offspring interactions, signaling alarm and maintaining flock-cohesion (Kroodsma et al., 1982). The most important uses of song appear to be courtship and

territory defense (Kroodsma and Byers, 1991), both of which are integral functions in the maintenance and continuity of the social structure of these species.

In many temperate songbird species, songs are limited to the male sex and females only produce call vocalizations. Zebra finches, *Taeniopygia guttata*, belong to the estrilidid group of finches and have been used extensively to study song learning. Zebra are highly gregarious and live in large social colonies but still most commonly learn the song of their father. This interplay between sociality and motivation (to learn from a specific tutor) make them an excellent system to study the role of social motivation in vocal learning.

We found that there is a temporal correlation between feeding-related behavior in both juveniles and fathers and adult male song from the father. Specifically, a larger proportion of fathers' responses to early begging were comprised of feeding with a small percentage of the response being song. As the juveniles matured and enter the sub-song phase of vocal development, an increasingly larger proportion of paternal response became adult male song. Additionally, this behavior had predictive value for song outcome. Fathers who had greater increases in their proportion of song responses to begging over the day 25-30 period of song development, raised juveniles whose own adult song was more like their fathers' song. We argue that this linking-at-source for food and song endows the father's song with motivational significance for the juvenile learner.

The motivational value of conspecific vocalization, specifically song, has long been appreciated. Both male and female chaffinches respond to the reinforcing effect of song, by choosing to perch on a particular perch and remaining on it in order to keep song playing (Stevenson-Hinde and Roper, 1975; Kling and Stevenson-Hinde, 1977; Reibel & Slater, 1998). Early studies also indicated that conditions that favor song development are also the ones that

promote maximum reinforcing effects of hearing that song in chaffiches (Stevenson-Hinde, 1972). Kling and Stevenson-Hinde, (1977) injected females who had been exposed to wild-type song in their first summer, with testosterone. They then tested these females' motivation to learn song by providing them with the ability to trigger normal song playback by a perching on a particular perch. These females learnt and produced song or at least subsong, indicating that birds will work to obtain song as a reinforcer.

Similarly, both male and female zebra finches show a preference for tutor (father) song in an operant conditioning task with song as reward (Riebel, Smallegange, Terpstra and Bolhuis, 2002). It is also known that behavior-contingent exposure to song acts as a reinforcer to young male finches with males preferring to sit on a perch that gives them exposure to song. Yoked controls that hear the same amount of song but not contingent on their own behavior, do not show this preference (ten Cate, 1991)

In human vocal learners, contingent vocal responsiveness leads to the learning of the specific acoustic parameters that are being reinforced (Goldstein and Schwade, 2008). In zebra finches, it is unclear if the father reinforces specific aspects of juvenile sub-song or masks their singing in counter singing bouts. Some recent data do indicate that counter-singing bouts in which the fathers' song follows the juveniles correlate positively with both sequential match and overall similarity score between juvenile and father songs. In an added dimension of social reinforcement, visual reinforcement from the mother in terms of fluff-ups and wing-strokes in response to juvenile song also hold predictive value for song outcome measures at crystallization (Otilia Menyhart, *unpublished data*). Thus it appears that song learning in zebra finches draws from distributed sources of social motivation. It is likely that initial rewarding interactions with the father reinforce the learning of a specific song and continued social feedback form both

parents guides the progression to crystalliation.

Overall discussion and concluding thoughts

Language, as defined here, is characterized by vocal learning. Vocal learning has been defined as "the ability to modify acoustic and/or syntactic structure of sounds produced, including imitation and improvisation". Vocal learning has been unequivocally attributed to 7 species in all, of which 3 are birds - songbirds, hummingbirds and parrots, and 4 are mammals - cetaceans, elephants, bats and humans (Jarvis, 2007). It is interesting to note that while the relative size of the isocortex in these organisms varies, all of them have a well-developed motivation/affect system. Janik and Slater (2000) differentiated vocal learning into production learning – ability to alter physical structure of sounds as a result of social interactions and contextual learning – the ability to use or comprehend a signal in a novel context. In organisms that display the relatively rare faculty of production language, social context seems to be necessary for this process. In organisms that don't specifically need to learn their species-typical vocalization (Owren et al., 1992; Seyfarth & Cheney, 1986; Winter et al., 1973), social context still modifies and modulates the nature and the content of such vocalizations (Roush & Snowdon, 1994, 1999; Seyfarth & Cheney, 1997).

The Evolution of Language within a Social Framework

The central argument here is that both the ontogeny and the phylogeny of language are constructed within the social realm because "motivation" is an essential part of linguistic behavior and the motivation to communicate only exists in a social context. One reason that has been offered for the motivated nature of social communication is the "parity" of language – that which is of significance for the speaker or the sender, is likely to also be of significance for the

hearer or receiver and this mutual significance is the incentive that drives social communication.

Vocal learning is significant because the overwhelming majority of species do not exhibit this capacity. Human vocal learning is significant because the complexity of its final product, language, is without precedent. This has lead many a scientist down the search for a neuroanatomical specialization for language in humans. Both Broca's area and the angular gyrus in the tempero-parietal-occipital junction have been upheld as trophy specializations, unique to humans at some point in the progression of the field (Deacon, 2007) and have subsequently been found to be present in the brains of other primates who don't have language (Deacon, 1992b; Ettlinger and Wilson, 1990). Similarly, voxel-based lesion-symptom mapping has indicated that neither Broca's nor Wernicke's areas are crucial to production or perception respectively, once the anterior insula or the middle temporal gyrus have been accounted for (Bates et al., 2003). Indeed, the left hemisphere appears to be involved in auditory processing in general. While brain activation in response to language does seem to be left lateralized, so called perisylvian "language areas" respond equally to linguistic and non-linguistic stimuli (Dick, et al., 2007) and lesions of Wernicke's area are more detrimental for nonverbal than verbal auditory processing (Saygin et al., 2003). Arguments for the linkage between the evolution of language and the disproportionate expansion of the human neocortex (Jerison, 1986; Deacon, 2003) can be reparteed by arguments against *unexpected* size changes in brain evolution (Semediferi et al., 2002; Finlay et al., 2005), the lack of any additional enlargement of language areas (parietal and temporal association cortex) in particular (Passingham, 1979), and increases in the sizes of the striatum, hippocampus, septum and amygdala (Stephan and Andy 1970, 1977) over evolutionary time. Similarly, evidence for the cortical basis of language (Aboitiz and Garcia, 1997) can be met by evidence for a) linguistic deficits caused by damage to regions that lie outside the cortex

(Geschwind, Quadfasel, and Segarra, 1968; Field et al., 2000), b) preservation of language in the face of cortical damage (Bellugi, 2000; Bates et al., 2004), and c) behavioural data that correlate language learning with motivation (Locke, 1993; Kuhl, 2007) These findings detract significantly from the idea of a cortically situated, self-sufficient language module (Chomsky, 1981). The cortex, seems to function best as a general-purpose learning device that probably plays a significant role in the complexity and memory component of modern *human* language but was not the solitary *a priori* driving force underlying the evolution of language. As illustrated in *Chapter 2*, the neocortical extent of the avian telencephalon is delimited by a relatively small region - the Wulst (Medina and Reiner, 2000). Since there is no indication that birds once had a neocortex that they used for the purposes of birdsong evolution and then discarded, it follows that a vocal communication system at least as complex as birdsong can develop in the absence of neocortical specializations targeted at language. To the extent that large parts of the avian cerebral hemispheres are derived from lateral and ventral pallia which give rise to limbic structures in the mammalian forebrain, which in turn are involved in affective and motivational processes, as well as form parts of the "social circuit" in the brain, we are again pointed in the direction of language being embedded within a social system of motivation. It isn't as much that we have special regions of the cortex devoted to language as it is that linguistic stimuli are "made special" to the brain via their links with social affect during development.

Kuhl et al. (2003, 2007) suggested two possible reasons as to why social interactions influence early speech learning in human ontogeny:

a) Social interaction motivates infants to learn through its properties of attention and arousal: Infants pay significantly higher attention to and are aroused more by a live person than an inanimate source and

b) The structured nature of social interactions provide infants with accessible information such as the relations between auditory labels, objects and speakers' intentions: "Attunement" to the communicative intentions of other humans in socially interactive situations that often include gaze following, joint focus of attention, shared referents, facilitate learning.

It is to be noted however, that these are not mutually exclusive mechanisms. Indeed, it seems reasonable to argue that the already structured information that is gathered from a socially guided process of learning is made more salient by the higher level of attention and arousal that is induced by the contingent and interactive nature of social interactions and this facilitates the learning of the said information, thus driving vocal learning.

Sources of structure

The ACCESS principles of infant vocal learning, as laid out by Goldstein et al. (2010a), predict that "infants should learn phonology, vocabulary, and syntax most effectively when the relevant structures are highlighted in their caregivers' speech by appearing contingently on the infants' behavior". An excellent example of the environment providing infants with structure that they can perceive and learn from can be found in infant-directed (ID) speech. Caregivers' speech to infants is often characterized by shorter utterances, longer pauses, higher pitch and wider pitch excursions (Fernald and Simon 1984, Fernald et al., 1989). The intonation of ID speech has been shown to effectively engage infant interest and is believed to be finely tuned to infant attention and arousal level (Fernald 1985). These data have led to the idea that ID speech is used by caregivers to facilitate infant arousal and motivation, which in turn facilitates infant learning of language. Specific features of ID speech make it particularly suited to the infant learner. ID speech makes the phonetic contrasts between both vowels and consonants more salient by exaggerating them and it has been suggested that this in turn makes words of the ambient

language easier to discriminate (Kuhl et al. 1997; Liu et al., 2007). Correspondingly, infants across cultures prefer to listen to ID speech (Cooper and Aslin, 1990; Hayashi, Tamekawa, & Kiritani, 2001), ID speech is correlated with enhanced activation in the temporal lobe of human infants (Bortfeld et al. 2007) and early maternal use of stretched acoustic cues in ID speech is predictive of later infant ability to hear speech distinctions (Liu et al., 2003)

Contingency:

My dissertation has emphasized the role of social contingency in vocal development in both humans and birds. Numerous studies with humans have highlighted the role of contingency in specific forms of reward-based learning. In adults, dorsal striatal (reward) activity is not reported in experiments where participants to not believe there is a connection between their actions and outcome (Breiter et al., 2001) and in infants, vocal learning is shaped by social contingency (Goldstein et al., 2003, 2008a, 2009a; Kuhl et al., 2003). Contingency seems to play a similar role in vocal learning in zebra finches. Adret (1993) used an operant conditioning paradigm wherein young male zebra finches were trained to press a button in order to hear playback of adult song. Yoked controls were also allowed to hear the same song, but were only passive participants and had no control over song presentation. It was found that the birds who could interactively control song presentation learnt the song while the yoked controls didn't. Narayan et al (2008) used a remote controlled car to simulate contingent and non-contingent (yoked control) responsiveness from a caregiver in humans. In the experimental condition, the car responded to infant vocalizations by approaching them every time they vocalized. The yoked controls received similar amount of stimulation from the car, but responding was not contingent on their vocalizations. The vocal quality of experimental, but not control, infants increased over the experimental period, similar to when infants interact with a contingent caregiver (Goldstein

et al. 2003). These data would indicate that it is interactive responsiveness that is contingent on the learner's behavior that drives learning and there is nothing "special about social".

However, in a similar experimental paradigm (as Adret, 1993), ten Cate (1991) found that experimental and control males do not differ in terms of how well they learn the song (ten Cate, 1991). Similarly, studies of filial imprinting have demonstrated that chicks, *Gallus gallus domesticus*, will press a lever to get exposure to a conspicuous object (Bateson and Reese, 1969) but do not become more strongly imprinted one the said object than controls (ten Cate, 1986). These data seem to indicate while responses contingent on behavior are reinforcing in and of themselves, the presence of a social conspecific recruits additional resources that facilitate learning. Like infant birds, infant humans learn from first exposure to phonetic contrasts in a foreign language provided the phonetic contrasts are provided to them via a socially interactive session with a conspecific. The extent of learning is determined by the amount of social interaction during the exposure period - overall attention and shared visual attention between the infant and the interactive social agent correlate with the degree to which individual infants learn foreign language contrasts (Kuhl et al. 2003, Conboy et al. 2008).

An idea that is central to this thesis is that the vocal learning that characterizes a species occurs in the context of their specific developmental niche. Learning in an ecological context arises from the interplay of environmental affordances and the learner's capacity to perceive such affordances (Gibson and Pick, 2000). Socially motivated influences on vocal learning illustrate the manner in which the developmental niche is structured. Within the developmental niche of an altricial species, the presence of adult caregivers and possibly other social conspecifics is an evolutionary expectation. The most likely source of contingent responsiveness for an infant learner over evolutionary time, was a social conspecific/caregiver. So the fact that

contingency is a mechanism of vocal learning does not detract from the role of social counterparts. Instead, it simply provides insight into the constitution of a social agent. Most, if not all social behaviors are typified by activation of motivation-related regions in the brain. The most parsimonious route to language then, becomes one that uses socially situated motivation to gate (aid and facilitate) language learning. Further, the developmental learning of species-typical vocal behavior is not indicative of an innate preference, but is instead indicative of structure (the presence of *language–experienced* conspecifics who will be motivated to interact with the young) that can be *expected* from the environment (Greenough and Black, 1992).

According to Dunbar's "gossip and grooming" hypothesis for the evolution of language (1996), language evolved in response to expansions in human groups that made physical grooming (as is commonly seen in most primate species) of all community members impossible. According to Falk's (2004) "motherese" theory of the evolution of language, language evolved from mothers' cooing vocalizations in response to infant distress at physical separation. The 18th century French naturalist Georges-Louis Leclerc, Comte de Buffon, suggested that origin of human language lay in the origin of the human family. Others (Deacon, 1997) expanded on this idea, bringing in the role of pair-bonding, which in turn permits infant learners to undergo the extended neotenous phase that characterizes altricial species (Gould, 1977). In essence, all of these theories embrace socialization and/or sociality-situated motivation as the critical driving force behind the evolution of language. As is illustrated through the role of a single receptor, Vasopressin, in mediating species-wide changes in social structure and affiliative behaviour (Young and Wang, 2004; Chapter 2), evolutionary mechanisms can act on specific aspects of an organism's niche to trigger a cascade of changes. It is within this context that this thesis proposes the evolution of a gated link between socio-motivational systems of the human forebrain and all-

purpose learning devices of cortical provenance (Chapter 2). And it is also within this context that this thesis argues for the fundamental role of social motivation in the ontogeny of vocal learning. Human vocal learning is a salience-tagged by the rewards of sociality. Throughout the phase of vocal learning, human infants attend preferentially to face stimuli over other classes of stimuli (Goren, 1975; Johnson 1991). They prefer human voices over other sounds (Eisenberg, 1976; Colombo and Bundy, 1981) and their mother's voices to other women's voices (DeCasper and Fifer, 1980). Touch is rewarding to both infants and mother (Field et al., 2004; Matthiesen et al., 2001). The infant activates the maternal brain almost exactly like cocaine (Strathearn et al., 2008). Early "proto-conversations" remain immersed in the motivational potency of these links. Within the domain of vocal interaction, every infant utterance and ever parent response is a consummately social act. And it is from within the motivation and structure embedded within this sociality that the remarkable feat of language emerges. Looks like social is special after all.

REFERENCES

Adolphs, R., (2008) Fear, faces, and the human amygdala. *Current Opinion in Neurobiology* 18:166–72.

Alcaro, A., Huber, R., and Panksepp, J., (2007) Behavioral functions of the mesolimbic dopaminergic system: an affective neuroethological perspective. *Brain Res Rev.* 6:283–321.

Bailey, J., and Meyerson, L. (1962) Vibrations as a reinforcer with a profoundly retarded child. *Journal of Applied Behavioral Analysis*, 2: 135-137.

Baptista,L.F., and Petrinovich,L., (1986) Song development in the white-crowned sparrow: social factors and sex differences, *Animal Behaviour*, 34(5) :1359-1371.

Barrera, M. E., and Maurer, D. (1981a). The perception of facial expressions by the three-month old. *Child Development*, 52: 203-206.

Bates, E. A. (2004). Explaining and interpreting deficits in language development across clinical groups: Where do we go from here? *Brain and Language, 88*(2): 248-253.

Bates, E., Wilson, S. M., Saygin, A. P., Dick, F., Sereno, M. I., Knight, R. T., Dronkers, N. F. (2003) Voxel-based lesion–symptom mapping. *Nature Neuroscience,* 6: 448-450.

Beecher, M. D., & Burt, J. M. (2004). The role of social interaction in bird song learning. *Current Directions in Psychological Science, 13,* 224-228.

Bellugi, U., Lichtenberger, L., Jones, W., Lai, Z., St. George, M. I. (2000). The neurocognitive profile of Williams syndrome: a complex pattern of strengths and weaknesses. *J. Cogn. Neurosci.*, 12(1): 7–29.

Berns, G.S., McClure, S.M., Pagnoni, G., Montague, P.R., (2001) Predictability modulates human brain response to reward. *J Neurosci* 21: 2793–2798

Bertenthal, B. I., Campos, J. J., and Barrett, K. C., (1984). Self-produced locomotion. In R. Emde & R. Harmon (Eds.), *Continuities and discontinuities in development*. New York: Plenum.

Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J. C. Goodman & H. C. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 167-224). Cambridge, MA: MIT Press.

Blood, A.J., Zatorre, R.J., Bermudez, P., and Evans, A.C., (1999) Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nat Neurosci* 2:382–387.

Bloom, K., Russell, A., & Wassenberg, K. (1987). Turn taking affects the quality of infant vocalizations. *Journal of Child Language, 14*, 211-227.

Boysson-Bardies, B., Halle, P., Sagart, L., & Durand, C. (1989). A cross-linguistic investigation of vowel formants in babbling. *Journal of Child Language, 16*, 1-17.

Boysson-Bardies, B., & Vihman, M. M. (1991). Adaptation to language: Evidence from babbling and first words in four languages. *Language, 67*, 297-320.

Boysson-Bardies, B. (1993). Ontogeny of language-specific syllabic productions. In B. Boysson-Bardies, S. Schonen, P. Jusczyk, P. MacNeilage, and J. Morton (Eds.), *Developmental neurocognition: Speech and face processing in the first year of life*. Dordrecht, The Netherlands: Kluwer.

Bradley, B.P., Mogg, K., and Millar, N.H. (2000) Covert and overt orienting of attention to emotional faces in anxiety. *Cogn Emot* 14(6): 789–808.

Campos, J.J., Hiatt, S., Ramsay, D., Henderson, C., and Svejda, M., (1978), The emergence of fear on the visual cliff. In M. Lewis & L, Rosenblum (Eds,), *The development of affect* (pp, 149-182), New York: Plenum.

Champagne, F., Diorio, J., Sharma, S., Meaney, M.J., (2001) Naturally occurring variations in maternal behavior in the rat are associated with differences in estrogen-inducible central oxytocin receptors. *Proc Natl Acad Sci* 98(22): 12736-41.

Champoux, M., and Suomi, S.J., (1994). Behavioral and adrenocortical responses of rhesus macaque mothers to infant separation in an unfamiliar environment. *Primates*, 35 (2): 191-202.

Clements, J. & Tracy, D. (1974a) Effects of tactile cueing and praise on undesirable classroom behavior of emotionally disturbed adoloscents. *The Training School Bulletin.*

Clements, J. & Tracy, D. (1974b) Effects of tactile reinforcement and verbal directions on attending behavior of young retarded children, *Journal of Developmental Disabilities.*

Clements, J. E.; Tracy, D. B. (1975) The Effects of Tactile Vs. Verbal Reinforcement on Attention to Task and Arithmetic Accuracy of Emotionally Disturbed Boys. Paper presented at the *Annual Meeting of the American Educational Research Association* (Washington, D.C., March 30-April 3, 1975)

Cochran, J.C., et al . 1992. Parsing attentional components during a simple reaction time task using sleep deprivation and amphetamine intervention. *Percept. Mot. Skills* 75 (3): 675–689.

Conboy, B. & Kuhl, P. K. (2007). Listening to language in culturally specific ways: Learning through social experience. In S. Braten (Ed.), *Foundations of Preverbal Intersubjectivity*: *Being Moved by Action-Perception, Music and Speech*. Cambridge University Press: Cambridge, England.

Conboy, B., Brooks, R., Taylor, R., Meltzoff, A., & Kuhl, P. (2008). Joint engagement with language tutors predicts brain and behavioral responses to second-language phonetic stimuli. Poster presented at the *XVIth International Conference on Infant Studies*, Vancouver, Canada.

Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech within the first month after birth. *Child Development*, 61: 1584–1595.

Delgado, M.R., Stenger, V.A., and Fiez, J.A. (2004). Motivation-dependent responses in the human caudate nucleus. *Cerebral Cortex*, 14:1022-1030.

Dick, F., Saygin, A. P., Gaspare, G., Sabrina, P., Simone, B., D'Amico, S., Wilson, S., Bates. E., Pizzamiglio, L. (2007). What is involved and what is necessary for complex linguistic and nonlinguistic auditory processing: evidence from functional magnetic resonance imaging and lesion data. *Journal of cognitive neuroscience* 19(5): 799-816.

Doupe, A. J. and Kuhl, P. K. (1999). Birdsong and human speech: common themes and mechanisms. *Ann. Rev. Neurosci.*, 22, 567-631.

Duchenne, Guillaume (1990). *The Mechanism of Human Facial Expression*. New York: Cambridge University Press. (Original work published 1862).

Dunbar, R., (1996) *Grooming, gossip, and the evolution of language*. London: Faber & Faber.

Elliott, R., Frith, C.D., and Dolan, R.J., (1997) Differential neural response to positive and negative feedback in planning and guessing tasks. *Neuropsychologia* 35:1395–404.

Ellis, N. C., (2006) Language Acquisition as Rational Contingency Learning, *Applied Linguistics,* 27: 1-24

Falk, D., Prelinguistic evolution in early hominins: Whence motherese? (2004) *Behavioral and Brain Sciences* 27: 491-503

Feldman,R., Weller,A., Zagoory-Sharon,O., and Levine,A., (2007), Evidence for a neuroendocrinological foundation of human affiliation: plasma oxytocin levels across pregnancy and the postpartum period predict mother–infant bonding, *Psychol. Sci.* 18, pp. 965–970.

Field, T., Hernandez-Reif, M., Diego, M., Feijo, L., Vera, Y., & Gil, K. (2004). Massage therapy by parents improves early growth and development. *Infant Behavior & Development*, 27, 435– 442.

Field, T. M., Woodson, R., Greenberg, R., and Cohen, D. (1982). Discrimination and imitation of facial expressions in neonates. *Science*, 218:179-181.

Finlay B.L., Cheung, D. and Darlington, R.B. (2005) Developmental constraints on or developmental structure in brain evolution? *In Attention and Performance XXI "Processes of Change in Brain and Cognitive Development*" Munakata, Y. and Johnson, M. Oxford University Press pp 131-162.

Fitzgerald,D.A., Angstadt,M., Jelsone,L.M., Nathan,P.J., Phan, K. L., (2006) Beyond threat: Amygdala reactivity across multiple expressions of facial affect, *NeuroImage*, 30(4): Pages 1441-1448

Francis, S., Rolls, E.T., Bowtell, R., McGlone, F., O'Doherty, J., Browning, A., Clare, S., and Smith, E., (1999) The representation of the pleasantness of touch in the human brain, and its relationship with taste and olfactory areas. *Neuroreport* 10:453–459.

Geschwind, N., Quadfasel, F. A. and Segarra, J. M. (1968). Isolation of speech area. *Neuropsychologia*, 6: 327-340.

Gibson, E, J., and Walk, R., (1960). The "visual cliff," *Scientific American*, 202: 64-71,

Goldstein, M. H., King, A. P., and West, M. J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proc. Nat. Acad. Sci.,* 100(13): 8030-8035.

Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science*.

Goldstein, M. H., & Schwade, J. A. (2009a) From birds to words: Perception of structure in social interactions guides vocal development and language learning. In M. S. Blumberg, J. H. Freeman, & S. R. Robinson (Eds.), *The Oxford Handbook of Developmental and Comparative Neuroscience*. Oxford University Press.

Goldstein, M. H., Schwade, J. A., & Bornstein, M. H. (2009b). The value of vocalizing: Five-month-old infants associate their own noncry vocalizations with responses from adults. *Child Development*, 80 (3), 636 – 644.

Goldstein, M. H., Waterfall, H., Lotem, A., Halpern, J., Schwade, J. A., Onnis, L., & Edelman, S. (2010a). General cognitive principles for learning structure in time and space. *Trends in Cognitive Sciences*, 14 (6), 249 - 258.

Goldstein, M. H., Schwade, J. A., Briesch, J., & Syal, S. (2010b). Learning while babbling: Prelinguistic object-directed vocalizations signal a readiness to learn. *Infancy*.

Gros-Louis, J. G., West, M. J., Goldstein, M. H., & King, A. P. (2006). Mothers provide differential feedback to infants' prelinguistic sounds. *International Journal of Behavioral Development*, 30 (6), 509 - 516.

Goren, C.C., Sarty, M., Wu PYK (1975) Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics* 56: 544–549.

Hansen, S., Bergval, A. H., Nyiredi, S., (1993). Interaction with pups enhances dopamine release in the ventral striatum of maternal rats: A microdialysis study. Pharmacol. *Biochem. Behav*. 45: 673–676.

Haugan, G.M. and McIntire, R.W., 1972. Comparisons of vocal imitation, tectile stimulation, and food as reinforcers for infant vocalizations. *Developmental Psychology* **6**, pp. 201–209.

Hayashi, A., Tamekawa, Y., & Kiritani, S. (2001). Developmental change in auditory preferences for speech stimuli in Japanese infants. *Journal of Speech, Language and Hearing Research*, 44, 1189–1200.

Jerison, H. J. (1976). The paleoneurology of language. *Ann. NY Acad. Sci*., 280:370-382.

Johnson, M.H., Dziurawiec, S., Ellis, H., and Morton, J., (1991) Newborn's preferential tracking of face-like stimuli and its subsequent decline. *Cognition* 40: 1–19.

Jusczyk, P. W. (1992). Developing phonological categories from the speech signal. In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 17-64). Timonium, MD: York Press.

Kamin, L. J. 1969. 'Predictability, surprise, attention, and conditioning' in B. A. Campbell and R. M. Church (eds): *Punishment and Aversive Behavior*. New York: Appleton-Century-Crofts, pp. 276–96.

Kensinger, E.A., Schachter, D.L., (2006). Processing emotional pictures and words: effects of valence and arousal. *Cogn. Affect. Behav. Neurosci.* 6: 110–126.

Kermoian, R., and Campos, J. J., (1988). Locomotor experience: A facilitator of spatial-cognitive development. *Child Development*, 59: 908-917.

King, A. P., West, M. J., & Goldstein, M. H. (2005). Nonvocal shaping of avian song development: Parallels to human speech development. *Ethology*, 111, 101- 117.

Koob , G.F., Nestler, E.J. ,(1997). The neurobiology of drug addiction. *J. Neuropsychiatry Clin. Neurosci*. 9 : 482–497. Links

Kuhl, P. K. (2007). Cracking the speech code: How infants learn language. *Acoust. Sci. and Tech*., 28:2-10.

Kuhl, P. K., Coffey-Corina, S., Padden, D. and Dawson, G. (2005). Links between social and

linguistic processing of speech in preschool children with autism: Behavioral and electrophysiological evidence. *Dev. Sci*., 8:1-12.

Kuhl, P. K., Tsao, F. M. and Liu, H. M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proc. Natl. Acad. Sci.,* 100:9096–9101.

Kuchuk, A., Vibbert, M., & Bornstein, M. H., (1986) The perception of smiling and its experiential correlates in three-moth-old infants. *Child Development*, 57. 1054-1061

La Barbera, J. D., Izard, C. E., Vietze, P., and Parisi, S. A. (1976). Four- and six-month-old infants' visual responses to joy, anger, and neutral expressions. *Child Development*, 47: 535-538.

Leshner , A.I., Koob, G.F. ,(1999). Drugs of abuse and the brain. *Proc. Assoc. Am. Physicians* 111 : 99–108.

Leibold,L.J., and Werner. L. A., Infant Auditory Sensitivity to Pure Tones and Frequency-Modulated Tones, *Infancy*, 12(2) August 2007 , pages 225 – 233

Lim, J., and Dinges, D.F. (2008), Sleep deprivation and vigilant attention. *Ann N Y Acad Sci.*, 1129: 305-22.

Locke J. (1993). *The Path to Spoken Language*. Harvard Univ. Press, Cambridge, MA.

Magill, R.A., Waters, W.F., Bray, G.A., Volaufova, J., Smith, S.R., Lieberman, H.R., McNevin, N., and Ryan, D.H. (2003) Effects of tyrosine, phentermine, caffeine D-amphetamine, and placebo on cognitive and motor performance deficits during sleep deprivation, *Nutr. Neurosci.* **6**: 237–246.

Marcus, G. F., Vijayan, S., Bandi Rao, S. & Vishton, P. M. Rule learning by seven-month-old infants. *Science* 283, 77-80 (1999).

Masataka, N. (2003). *The onset of language*. Cambridge: Cambridge University Press.

Matthiesen, A.S., Ransjo-Arvidson A.B., Nissen E., Uvnas-Moberg, K. (2001). Postpartum

maternal oxytocin release by newborns: Effects of infant hand massage and sucking. *Birth*, 28, 13–19.

Maye, J., Weiss, D. J., & Aslin, R. N. (2008). Statistical phonetic learning in infants: Facilitation and feature generalization. *Developmental Science*, 11, 122–134.

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition,* 82, B101–B111.

McCandliss, B. D., Fiez, J. A., Protopapas, A., Conway, M., & McClelland, J. L. (2002). Success and failure in teaching the r-l contrast to Japanese adults: predictions of a hebbian model of plasticity and stabilization in spoken language perception. *Cognitive, Affective, and Behavioral Neuroscience*, 2(2), 89-108.

Medina, L., and Reiner, A. (1994). Distribution of choline acetyltransferase immunoreactivity in the pigeon brain. *J. Comp. Neurol.*, 342(4):497-537.

Messinger, D. S., Fogel, A., & Dickson, K. L. (2001). All smiles are positive, but some smiles are more positive than others. *Developmental Psychology*, 375, 642-653.

Mogg, K., and Bradley, B.P. (2002) Selective orienting of attention to masked threat faces in social anxiety. *Behav Res Ther* 40: 1403–1414.

O'Doherty, J., Kringelbach, M.L., Rolls, E.T., Hornak, J., and Andrews, C., (2001a) Abstract reward and punishment representations in the human orbitofrontal cortex. *Nat Neurosci* 4:95–102

Oller, D.K., (2000) *The emergence of the speech capacity*. Laurence Erlbaum Associates, Publishers, Mahwah, New Jersey, London.

Pedersen, C.A., Caldwell, J.D., Walker, C., Ayers, G., Mason, G.A., (1994) Oxytocin activates the postpartum onset of maternal behavior in the ventral tegmental and medial preoptic areas. *Behav Neurosci* 108:1163–1171

Pederson, C.A., (1997) Oxytocin Control of Maternal Behavior Regulation by Sex Steroids and Offspring Stimuli, Issue: Integrative Neurobiology of Affiliation, The. *Annals of the New York Academy of Sciences,* (807) 126-145.

Pedersen, C. A., Caldwell, J. D., Walker, C., Ayer, G., Mason, G. A., (1994). Oxytocin activates the postpartum onset of rat maternal behavior in the ventral tegmental and medial preoptic areas. *Behav. Neurosci*. 108: 1163–117

Poulson, C.L.,(1983) Differential reinforcement of other-than-vocalization as a control procedure in the conditioning of infant vocalization rate, *Journal of Experimental Child Psychology*, 36 (3): 471-489

Redouté, J., Stoléru, S., Grégoire, M.C., Costes,N., Cinotti, L., Lavenne, F., Le Bars, D., Forest, M.G., and Pujol, J.F., (2000) Brain processing of visual sexual stimuli in human males. *Hum Brain Mapp* 11:162–177.

Rheingold, H. L., Gerwitz, J. L., & Ross, H. W. (1959). Social conditioning of vocalizations in the infant. *Journal of Comparative and Physiological Psychology, 52*, 68-73.

Rollins, P. R. (2003). Caregivers' contingent comments to 9-month-old infants: Relationships with later language. *Applied Psycholinguistics, 24*, 221-234.

Rolls, E.T., (1999) *The Brain and Emotion*. Oxford: Oxford University Press

Rolls, E.T., Burton, M.J., and Mora, F., (1980) *Brain Res* 194: 339–357

Rolls, E.T., O'Doherty, J., and Kringelbach, M.L., (2003) Representations of pleasant and painful touch in the human orbitofrontal and cingulate cortices. *Cerebral Cortex* 13:308–17

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science,* 274, 1926-1928.

Sander, K., Frome, Y., Scheich, H., (2007) FMRI activations of amygdala, cingulate cortex, and auditory cortex by infant laughing and crying. *Hum Brain Mapp* 28:1007–1022.

Saygin, A. P., Dick, F., Wilson, S. W., Dronkers, N. F., & Bates, E. (2003). Shared neural resources for processing language and environmental sounds: Evidence from aphasia. *Brain*, 126, 928–945.

Schooler, L. J. and Anderson, J. R. (1997) 'The role of process in the rational analysis of memory,' *Cognitive Psychology 32*(3): 219–50.

Schultz, W., Predictive reward signal of dopamine neurons. *J. Neurophysiol*. 80: 1–27, 1998

Schultz W. The Reward Signal of Midbrain Dopamine Neurons (1999), *News Physiol Sci*.14: 249-255.

Semendeferi, K., Lu, A., Schenker, N., and Damasio, H. (2002). Humans and great apes share a large frontal cortex. *Nat. Neurosci*., 5: 272 – 276.

Shahrokh, D. K., Zhang, T., Diorio, J., Gratton, A., and Meaney, M. J. (2010) Oxytocin-Dopamine Interactions Mediate Variations in Maternal Behavior in the Rat. *Endocrinology*, 151(5): 2276-2286.

Smotherman, W. P., Bell, R. W., Starzec, J., Elias, J., Zachman, T. A., (1974). Maternal responses to infant vocalizations and olfactory cues in rats and mice. *Behav. Biol.* 12, 55–66.

Soussignan, R. (2002). Duchenne Smile, emotional experience, and automatic reactivity: A test of the facial feedback hypothesis. *Emotion*, 2(1): 52-74.

Stephan, H., and Andy, O. J. (1977). Quantitative comparison of the. amygdala in insectivores and primates. *Acta. Anat.,* 98:130-153.

Stevenson-Hinde,J., and Roper,R.,(1975) Individual differences in reinforcing effects of song, *Animal Behaviour*, 23(4) :729-734

Stock, S., and Uvnas-Moberg,K., Increased plasma levels of oxytocin in response to afferent electrical stimulation of the sciatic and vagal nerves and in response to touch and pinch in anaesthetized rats, *Acta Physiol.* Scand. 132 (1988), pp. 29–34.

Stoel-Gammon, C. (1992). Prelinguistic vocal development: Measurement and predictions. In C. A. Ferguson, L. Menn, and C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 439-456). Timonium, MD: York Press.

Strathearn,L., Li,J., Fonagy,P., Montague, P. R.,(July 2008) What's in a Smile? Maternal Brain Responses to Infant Facial Cues. *Pediatrics* 122(1): pp. 40-51

Sullivan, R.M., and Hall, W.G., (1988) Reinforcers in infancy: classical conditioning using stroking or intra-oral infusions of milk as UCS *Dev Psychobiol*. 21(3):215-23.

Sullivan,R.M., Taborsky-Barba, S., Mendoza,R., Itano, A., Leon, M., Cotman,C.W., Payne, T.F., and Lott,I., (1991) Olfactory Classical Conditioning in Neonates 87(4): 511–518.

Tamis-LeMonda, C. S., Bornstein, M. H., & Baumwell, L. (2001). Maternal responsiveness and children's achievement of language milestones. *Child Development, 72*(3), 748-767.

Tricomi, E.M., Delgado, M.R., and Fiez, J.A. (2004). Modulation of caudate activity by action contingency. *Neuron*, 41:281-292.

Tricomi, E., Delgado, M. R., McCandliss, B. D., McClelland, J. L. and Fiez, J. A. (2006).Performance feedback drives caudate activation in a phonological learning task. *J.Cognit. Neurosci.* 18:1029-1043.

Weisberg, P. (1963). Social and nonsocial conditioning of infant vocalizations. *Child Development, 34*, 377-388.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development, 7*, 49-63.

Werker, J. F., Pons, F., Dietrich, C., Kajikawa, S., Fais, L., Amano, S. (2007) Infant-directed speech supports phonetic category learning in English and Japanese, *Cognition*, Volume 103(1):147-162

West, M. J., and King, A. P. (1988). Female visual displays affect the development of male song in the cowbird. *Nature*, 334(6179):244-246.

Whalen, D. H., Levitt, A., & Wang, Q. (1991). Intonational differences between the reduplicative babbling of French- and English-learning infants. *Journal of Child Language, 18*, 501-516.

Whalen, P.J., Kagan, J., Cook, R.G., Davis, F.C., Kim, H., Polis, S., McLaren, D.G., Somerville, L.H., McLean, A.A., Maxwell, J.S., and Johnstone, T., (2004). Human amygdala responsivity to masked fearful eye whites. *Science* 306: 2061.

Yang, T.T., Menon, V., Eliez, S., Blasey,C., White, C.D., Reid, A.J., Gotlib, I.H., and Reiss, A.L., (2002) Amygdalar activation associated with positive and negative facial expressions, *NeuroReport* 13: pp. 1737–1741

Young, L. J. , Lim, M. M. , Gingrich, B. and Insel, T. R. Cellular mechanisms of social attachment. *Horm. Behav*. 40, 133−138

Young, L. J., and Wang, Z. (2004). The neurobiology of pair bonding. *Nature neuroscience, 7*(10), 1048-1054.

Young-Browne, G., Rosenfeld, H. M., & Horowitz, F. D. (1977). Infant discrimination of facial expressions. *Child Development*, 48, 555-562.

Yoshida, K. A., Pons, F., Maye, J., & Werker, J. F. (2010). Distributional phonetic learning at 10 months of age. *Infancy*, 15(4), 420-433.